أثر معالجة البيانات المفقودة وحجم العينة ونسبة الفقد على قوة الاختبار الإحصائي والدلالة العملية

The impact of handling missing data, sample size, and missingness ratio on the power of statistical tests and practical significance

إعداد

سعد بن عوض بن علي العمري

Saad Awad Ali Al-Omari

قسم علم النفس -كلية التربية جامعة أم القرى

أ.د/ ياسر بن عبدالله حفني

Prof. Yasser Abdullah Hafni

أستاذ القياس والتقويم والإحصاء بكلية التربية جامعة أم القرى

Doi: 10.21608/jasep.2025.423372

استلام البحث: ٢٢ / ٢ /٢٠٢٥

قبول النشـــر: ١٨ / ٣ / ٢٠٢٥

العمري، سعد بن عوض بن علي وحفني، ياسر بن عبدالله (٢٠٢٥). أثر معالجة البيانات المفقودة وحجم العينة ونسبة الفقد على قوة الاختبار الإحصائي والدلالة العملية. ***المجلة العربية للعلوم التربوية والنفسية***، المؤسسة العربية للتربية والعلوم والآداب، مصر،٩(٤٨)، ٢٥٣ – ٣١٠.

http://jasep.journals.ekb.eg

أثر معالجة البيانات المفقودة وحجم العينة ونسبة الفقد على قوة الاختبار الإحصائي والدلالة العملية

المستخلص:

هدفت الدّراسة إلى دراسة تأثير طرق معالجة البيانات المفقودة على قوة الاختبار الإحصائي والدلالة العملية، من خلال المقارنة بين ثلاثة طرق (تعظيم التوقعات، التعويض بالانحدار، التعويض المتعددة) لتقدير البيانات المفقودة ولتحقيق هدف الدراسة تم استخدام بيانات افتراضية تحاكي الظروف التجريبية المتمثلة في نسبة القيم المفقودة ولها ثلاث مستويات (١٠%، ٢٠%، ٤٠%) وباستخدام أحجام عيّنات مختلفة (٢٠٠، ٦٠٠، ١٠٠٠) على مفحوصين افتراضيين واختبار مكوّن من (٤٠) فقرة تم توليدها من خلال برنامج (WINGEN3)، وفق أسلوب المحاكاة لتوليد أرقام عشوائية حيث تم بناء ثلاثة نماذج إحصائية لمحاكاة أسلوب (اختبار ت لعينة واحدة، اختبار ت لعينتين مستقلتين، واختبار تحليل التباين الأحادي لثلاث عينات) بحيث تتوافر فيها الافتراضات الخاصّة بها، وتمّ حساب قوة الاختبار الإحصائي للحالات المذكورة، ثم تمّ عمل فقد لهذه البيانات باستخدام برنامج (R) بنسب فقدٍ بلغت (١٠%، ٢٠%، ٤٠%) على التوالي وفق آلية الفقد (MCAR) وتحت نمط فقد (النّمط التام) وبعد ذلك تمّ تقدير القيم المفقودة باستخدام الطرق الثلاثة كلٍ على حِدة، وإجراء الاختبارات الإحصائية الثلاثة لها وحساب قيمة قوة الاختبار الإحصائي والدلالة العملية، للحالات المذكورة، وتمت المفاضلة بين الطرق الثلاث عند كل أسلوب إحصائي بمقارنة قوة الاختبار الإحصائي والدلالة العملية لتلك النماذج مع قوة الاختبار الاحصائي والدلالة العملية للنموذج الإحصائي التام (بدون فقد)، كما تم الكشف عن الدلالة الإحصائية عند مستوى الدلالة ($\alpha \leq$ ٠,٠٥) للفروق بين المتوسطات الحسابية لقوة الاختبار الإحصائي بدلالة قيمة مربع ايتا التي تُعزى لأثر طريقة المعالجة ونسبة الفقد حيث كانت لصالح طريقة تعظيم التوقعات والتعويض المتعدد مقابل طريقة التعويض بالانحدار كما تبين وجود فروق لنسب الفقد لصالح الأقل (١٠%) مقابل (٤٠%)، كما تبين عدم تأثر الدلالة العملية لكافة المجموعات ولطرق المعالجة الثلاثة وكذلك مع أحجام العينات المختلفة وبمقارنتها مع النموذج التام. وتوصي الدراسة باستخدام طريقة تعظيم التوقعات والتعويض المتعدد بالدراسات التربوية والنفسية في معالجة البيانات المفقودة، بعد التأكد من استيفاء افتراضات استخدام طرق التعويض قبل الاعتماد عليها.

الكلمات المفتاحية: التعويض بـ تعظيم التوقعات، التعويض المتعددة، التعويض بـالانحدار، البيانات المفقودة، نسبة الفقد، حجم العينة، قوة الاختبار الإحصائي، الدلالة العملية.

ISSN: 2537-0464 eISSN: 2537-0472

Abstract:

The study aims to study the impact of missing data handling methods on the power of the statistical test and practical significance by comparing three methods: expectation-maximization (EM), regression compensation, and multiple compensation to estimate missing data. To achieve the aim of the study, the researcher used hypothetical data that simulates the experimental conditions represented by the percentage of missing values, which has three Levels (10%, 20%, 40%) and using different sample sizes (200, 600, 1000) on virtual examinees and a test consisting of (40) items generated through WINGEN3 program according to the simulation method to generate random numbers, where three statistical models were built for simulating the method (T Test for one sample, T Tests for two independent samples, and a one-sample analysis of variance test for three samples) so that their assumptions are met, and the power of the statistical test for the cases mentioned was calculated, then loss of this data was done using R program in proportions, they reached (10%, 20%, and 40%) respectively according to the missing data mechanism (Missing Completely At Random/ MCAR) and under a missing mode (Complete Mode). After that, the missing values were estimated using the three methods separately, and the three statistical tests were performed for them and the value of the power of the statistical test and the practical significance for the mentioned cases were calculated. A comparison was made between the three methods for each statistical method by comparing the power of the statistical test and the practical significance of those models with the power of the statistical test and the practical significance of the complete statistical model (without loss). Also, the statistical significance was revealed at the level of significance ($0.05 \geq \alpha$) for the differences between the arithmetic means of the power of the statistical test in terms of the value of the Eta square, which

ISSN: 2537-0464 eISSN: 2537-0472

is attributed to the impact of the handling method and the loss percentage, where it was in favor of the expectation maximization and multiple compensation methods versus the regression compensation method. It also showed that there were differences in the loss percentages in favor of the least (10%) versus (40%). It was also shown that the practical significance was not affected for all groups and the three treatment method as well as with different sample sizes and comparing them with the complete model. The study recommends using the expectation-maximization and multiple compensation methods in educational and psychological studies after ensuring that the assumptions for using compensation methods are satisfied before relying on them in handling missing data.

Key Words: Compensation by expectation-maximization - multiple compensation, regression compensation, missing data, missing data ratio, sample size, power of statistical test, practical significance**.**

مقدمة:

تعتبر البيانات في العلوم البحثية بمثابة العامود الفقري لها، والأساس الذي يعتمد عليه الباحثون في الإجابة عن تساؤلاتهم ولاختبار فرضياتهم البحثية. وحيث شهدت البحوث والدراسات بمجال العلوم الإنسانية منذ مطلع القرن الحادي والعشرين، ظهور العديد من الأساليب والمعالجات الإحصائية الحديثة مستفيدة من تطبيقات الذكاء الاصطناعي، والاهتمام بنمذجة البيانات الإحصائية وفق أسلوب المحاكاة، وظهور البرامج الإحصائية المتخصصة، جميعها عوامل ساهمت في تجويد وتحسين أساليب التحليل الإحصائي المستخدمة بالدّراسات التجريبية والميدانية، التي يقومون بإجرائها وهناك تقدم ملموس للأساليب والإجراءات الإحصائية التي تتعامل مع المشاكل والتحيزات الناتجة عن البيانات المفقودة، ولذلك فإن مشكلة فقدان البيانات هي مشكلة شائعة في العديد من المجالات، بما في ذلك التعليم وعلم النفس والعلوم الاجتماعية وغيرها. وهذا يستدعي دراسة مشكلة البيانات المفقودة وتقديم الطريقة المناسبة للتعامل معها، للحصول على نتائج دقيقة للبحوث.

ومع تزايد الاهتمام بالجودة في الأبحاث العلمية، أصبح التعامل مع البيانات المفقودة ومعالجتها موضوعًا ذا أهمية. خصوصًا مع تقدم البرامج الإحصائية التي

ISSN: 2537-0464 eISSN: 2537-0472

تتعامل مع البيانات البحثية، حيث أصبحت هناك العديد من الأساليب والطرق المتاحة للتعامل مع البيانات المفقودة. هذه الأساليب تتطلب فحصًا ومقارنة للوصول إلى نتائج علمية تساعد الباحثين على اختيار الطريقة المناسبة للتعامل مع البيانات المفقودة وكيفية استخدامها لتحسين نتائج الأساليب الإحصائية. هذا بالمقارنة مع الأساليب الإحصائية التي لا تتضمن التعويض عن البيانات المفقودة. لتخدم الهدف النهائي في الدراسة وهو الوصول إلى استنتاجات صحيحة يمكن تعميمها حول الظاهرة المدروسة، من خلال تحسين النتائج وتحقيق التكامل بين البيانات المجمعة وعملية التعويض باستخدام البيانات المتاحة (بني عواد والمومني، ٢٠١٨).

فالبعض يلجأ للتخلص من أي فرد قام بعملية الاستجابة وله قيم مفقودة لأي من الفقرات، وتحليل الأفراد الذين استجابوا لجميع الفقرات فقط، وهذا يعني تخفيض في حجم العينة مما يؤثر سلباً في القوة الإحصائية (Pigott , 2001).

ولقد عمد المتخصصون في قضايا القياس النفسي والتربوي إلى الكشف عن العديد من الطرق لمساعدة الباحثين في التعامل مع مشكلة البيانات المفقودة، وذلك لضمان الوصول إلى نتائج موثوقة تقود إلى اتخاذ قرارات صحيحة في رفض الفرض الصفري وقبول الفرض البديل. والذي يُعرف بقوة الاختبار الإحصائي (Power of Statistical test)، والذي يُقصد به الاحتمالية في رفض الفرضية الصفرية عندما تكون خاطئة، ويُرمز له بالرمز (١- بيتا) (علام، ٢٠٠٣)، وعادة ما يتم قياسه من خلال حجم الأثر، الذي يُعتبر بدوره من المؤشرات العملية التي تعطي نتائج البحوث قيمة عملية للأخذ بنتائجها، والذي يُعتبر من الدلائل المهمة إلى جانب الدلالة الإحصائية لمعرفة قوة الاختبار الإحصائي (علام، ٢٠١٠؛ أبو بدر، ٢٠٢٠). لذلك من الضروري النظر بعناية في نسبة البيانات المفقودة وتأثيرها المحتمل على صحة الاستنتاجات المستخلصة من الدراسة .كما يعد حجم العينة عاملاً مهمًا يجب مراعاته عند تقييم قوة الاختبار الإحصائي والدلالة العملية. حيث يوفر حجم العينة المزيد من القوة لاكتشاف التأثيرات ذات الأهمية الإحصائية، كما ينبغي التأكد من أن حجم العينة مناسب للهدف البحثي والتحليل الإحصائي للدراسة، فضلاً عن كون طريقة المعالجة للبيانات المفقودة، ونسبة الفقد، وحجم العينة، كلها عوامل تؤثر على قوة الاختبار الإحصائي والدلالة العملية. ومن الأهمية بمكان النظر بعناية في هذه العوامل عند إجراء البحوث لضمان سلامة وموثوقية نتائج الدراسة.

ويعد الذكاء الاصطناعي والذي يعتبر طريقة جذابة للتعامل مع القيم المفقودة. ورغم تصميمه المبكر والعديد من المزايا التي يقدمها مقارنة بالأساليب والطرق التقليدية، لا يزال هناك تطبيق محدود لخوارزمية تعظيم التوقعات-Expectation) Maximization Algorithm (EM)). (الحيالي، ٢٠١٣)، وتتميز طريقة تعظيم التوقعات بالعديد من المزايا التي تتغلب على الأساليب التقليدية، حيث يتم استبدال

ISSN: 2537-0464 eISSN: 2537-0472

القيمة المتوقعة بناءً على المتغيرات المتاحة لكل حالة بالبيانات المفقودة. كما تتعامل مع مشكلة التعويض الفردي الذي يتجاهل الاختلافات المحتملة بين الأفراد، وتجنب المبالغة في تقدير مستوى الدقة (Graham,2009).

مشكلة الدراسة وتساؤلاتها:

تعد البيانات المفقودة تحديًا في البحث والتحليل، حيث يصعب حساب القيم المفقودة باستخدام غالبية الأساليب الإحصائية. على الرغم من اهتمام الباحثين بفهم هذه القيم، إلا أن التعامل معها يظل أمرًا صعبًا من الناحية النظرية والحسابية، يلجأ الباحثون إلى استخدام طرق وأدوات غير صحيحة للتعامل مع هذه القيم المفقودة حيث وصفها دي جل (de Gil, 2015) بأنها الأساليب غير الرسمية والمرتجلة لملء البيانات لتؤدي إلى زيادة التحيز وتقليل موثوقية النتائج، لذا يوجد حاجة ماسة لفهم الطرق المتنوعة لحساب القيم المفقودة في التحليل وكيفية تطبيقها في الدراسات النفسية والتربوي، كما أوضح بول أليسون (Allison, 2009) من الصعب جداً تحليل البيانات التي لا تتوفر لدينا، وهذا يفتح الباب أمام الاستفادة بالخبراء والمختصين في المجال لتطبيق أفضل الطرق والأساليب حول كيفية التعامل مع البيانات المفقودة.

وقد أشار بوري (Bori, 2013) إلى أن البيانات المفقودة تؤثر سلباً على دقة الأساليب الإحصائية وتضعف قوة الاختبار الإحصائي، هذا لأن كل أسلوب إحصائي يعتمد على افتراضات معينة ومعلومات كاملة عن جميع المتغيرات المدرجة في التحليل. ويمكن لعدد صغير نسبياً من البيانات المفقودة التأثير على بعض المتغيرات حيث أنه يقلل بشكل كبير من حجم العينة، مما يؤثر على دقة فترات الثقة، ويضعف القوة الإحصائية، ويجعل تقدير المعلمة متحيزاً.

ولقد قام (Karahalios et al., 2012) بتنفيذ استعراض منهجي لكيفية تقديم التقارير والتعامل مع البيانات المفقودة في الأبحاث العلمية حيث تم مراجعة ٨٢ دراسة، وأعلن عن أن ٦٦ منها ما يمثل (٨٠%) تعبر عن وجود بيانات مفقودة بشكل عام. وأفاد أن ١٣ منها (١٦%) تعبر عن الأسباب التي أدت إلى فقدان البيانات، ولم يذكر أن ١٤ منها ما يمثل (١٦%) المنهجية الإحصائية المستخدمة للتعامل مع البيانات المفقودة، أشار الباحثون إلى أن ٣٨ من هذه الدراسات تم استبعادها، واستخدمت (٦٦%) من الدراسات طريقة التحليل الكامل للحالة.

فضلا عما ذكر فقد تتعدد وتتعقد الطرق الإحصائية المستخدمة للتعامل مع البيانات المفقودة، مما يزيد من تعقيد مشكلة البيانات المفقودة. ولقد بذل الإحصائيون جهودًا لتطوير ومعرفة أفضل الطرق لتقدير القيم المفقودة في مختلف بيانات البحث العلمي. كل طريقة من هذه الطرق لها شروطها وافتراضاتها الخاصة، وفقًا لنمط الفقد

ISSN: 2537-0464 eISSN: 2537-0472

وآليته (Mcknight, Sidani, Figueredo & Mcknight, 2007). ولكن الخطأ الشائع الذي يقع فيه الإحصائيين هو التعامل مع القيم المفقودة دون الأخذ في الاعتبار أنماطها والأساليب الإحصائية التي تحدث فيها ونسبة الفقد وحجم العينة، هذا يؤدي إلى زيادة تكرار الأخطاء الإجرائية في التحليل الإحصائي، وبالرغم من التقدم الهائل في معالجة مشكلة البيانات المفقودة، وخاصة في تطوير طرق معالجتها وتقديرها، إلا أن الطرق التقليدية التي تستند إلى الحذف ما زالت مستخدمة بشكل واسع، السبب في ذلك هو سهولة استخدام هذه الطرق، بينما الطرق الحديثة تتطلب وقتاً أطول ومعرفة أعمق لكي يتم فهمها وتطبيقها بشكل صحيح بدلاً من الطرق التقليدية (Enders,2010).

ومن خلال العرض السابق لمشكله الدراسة يتضح أن هناك ضعف في الجانب التجريبي المرتبط بالقيم المفقودة، بالإضافة إلى اتساع الفجوة الناتجة من ضعف وعي الدارسين بأهمية التعامل مع القيم المفقودة واسهاماته في التقليل من الأخطاء التي يقع بها الكثير من الباحثين في معالجاتهم الإحصائية والتي من الممكن ان تساهم في الوصول الى الحل العلمي النظري، لتساعد الباحث في ظل تحقق الافتراضات المعيارية لاختيار الطريقة المُثلى الأقل تأثيراً على قوة الاختبار الإحصائي وذات الدلالة العملية الأعلى، وفي إعطاء تفسيرات دقيقة.

وبذلك يمكن تحديد مشكلة الدراسة في التساؤلات التالية:

١. ما تأثير طرق معالجة البيانات المفقودة (تعظيم التوقعات، الانحدار، التعويض المتعدد) ونسبة الفقد وحجم العينة على قوة الاختبار الإحصائي والدلالة العملية لمجموعة واحدة؟
٢. ما تأثير طرق معالجة البيانات المفقودة (تعظيم التوقعات، الانحدار، التعويض المتعدد) ونسبة الفقد وحجم العينة على قوة الاختبار الإحصائي والدلالة العملية لمجموعتين؟
٣. ما تأثير طرق معالجة البيانات المفقودة (التعويض المتعدد، الانحدار، أقصى توقع) ونسبة الفقد وحجم العينة على قوة الاختبار الإحصائي والدلالة العملية لثلاث مجموعات؟

أهداف الدراسة:

هدفت الدّراسة إلى استقصاء تأثير طرق معالجة البيانات المفقودة على قوة الاختبار الإحصائي والدلالة العملية " دراسة محاكاة"، من خلال تحقيق الأهداف التالية:

١. الكشف عن تأثير طرق معالجة البيانات المفقودة (تعظيم التوقعات، الانحدار، التعويض المتعدد) ونسبة الفقد وحجم العينة على قوة الاختبار الإحصائي والدلالة العملية لمجموعة واحدة.

ISSN: 2537-0464 eISSN: 2537-0472

٢. الكشف عن تأثير طرق معالجة البيانات المفقودة (تعظيم التوقعات، الانحدار، التعويض المتعدد) ونسبة الفقد وحجم العينة على قوة الاختبار الإحصائي والدلالة العملية لمجموعتين.
٣. الكشف عن تأثير طرق معالجة البيانات المفقودة (تعظيم التوقعات، الانحدار، التعويض المتعدد) ونسبة الفقد وحجم العينة على قوة الاختبار الإحصائي والدلالة العملية لثلاث مجموعات.

أهمية الدراسة:

تكمن أهمية الدراسة من خلال تسليط الضوء على عدد من طرق معالجة البيانات المفقودة قليلة الاستخدام من قبل الباحثين وهي (تعظيم التوقعات، الانحدار، التعويض المتعدد) في البحوث النفسية والتربوية. وحول تأثير معالجة البيانات المفقودة على فعالية بعض الأساليب الإحصائية المعلمية (ت، ف) بمحاكاة أحجام عينات مختلفة ونسبة فقد مختلفة ومعرفة مدى تأثيرها ذلك على قوة الاختبار الإحصائي والدلالة العملية.

حيث تبرز الأهمية للدراسة في جانبين:

- **الأهمية النظرية وتتمثل بالجوانب التالية:**

– المساهمة في الأدب النفسي المتعلق بأساليب وطرق فقد البيانات في ضوء ظروف افتراضية لمحاكاة البيانات الحقيقية، حيث يتوقع أن تكشف نتائج هذه الدراسة عن أكثر الطرق كفاءة وفاعلية، مع التأطير المفاهيم لطرق معالجة البيانات المفقودة وبيان أنماط وآليات فقد البيانات كذلك مساهمتها في التراكم المعرفي بأساليب معالجة فقد البيانات في ضوء ظروف وافتراضات معيارية لمعالجة البيانات المفقودة.

– المساهمة في الأدبيات التعليمية حول كيفية دراسة وتقييم تقنيات واستراتيجيات البيانات المفقودة في بيئة افتراضية تعتمد على أنماط البيانات الافتراضية التي تحاكي البيانات المفقودة في الحياة الواقعية.

– من المتوقع أن تكشف نتائج هذه الدراسة عن الأساليب الأكثر جدوى وفعالية، والسياق النظري لمعالجة البيانات المفقودة، وتوضيح أنماط وآليات الخسارة، بالإضافة إلى فوائدها في توليد المعرفة فيما يتعلق بتقنيات البيانات المفقودة عند العمل في ظل الظروف العادية.

– تأتي الدراسة الحالية استجابة نظرية للإسهام في حركة القياس النفسي والاختبارات لتحديد الطَّريقة الأفضل لمعالجة البيانات المفقودة، في ضوء اختلاف أحجام العينات واختلاف نسب الفقد واختلاف المعالجات والأساليب الإحصائية التي يتم من خلالها البحث والاستقصاء بأثر أسلوب معالجة البيانات على قوة الاختبار الإحصائي والدلالة العملية بناءً على الظُّروف المحيطة بعمليّة التحليل وشكلها، والَّتي

ISSN: 2537-0464 eISSN: 2537-0472

يعتقد الباحث بأهميّتها في عمليّة المقارنة بين طرق معالجة البيانات المفقودة الشائعة والحديثة حيث تزوّد الأدب البحثي بنتائج متعلّقة بأهم الظُروف المؤثّرة على دقة التحليل من حيث قوة الاختيار الإحصائي والدلالة العملية.

- نظراً لقلة الدراسات الميدانية بالمجال – على حدود علم الباحث والتي - تتناول طرق حديثة لمعالجة البيانات المفقودة وخاصّة أنها طرق واعدة من حيث تطويرها وتكييفها لتقنيات الذكاء الاصطناعي بالمستقبل تحت مجموعة تطبيقات التعلّم الآلي مما قد يُسهم بفتح المجال أمام الباحثين لإجراء المزيد من الدراسات النظرية.

- **الأهمية العملية وتتمثل بالجوانب التالية:**

من خلال نتائج هذه الدراسة

- يمكن أن يستفيد الباحثون في التغلب على مشكلة فقد البيانات من خلال معرفة أكثر الطرق ملائمة لاستخدامها في دراساتهم، كما أنها تبحث في صحة الممارسة التي يتّبعها الباحثون.

- يمكن أن تعتبر نتائج الدراسة مطورو المقاييس والاختبارات في معرفة أكثر الطرق ملائمة في التعامل مع البيانات المفقودة، وأيضاً في معرفة ما هي الطريقة المناسبة والتي يمكن ألا تتأثر بالمعالجات المستخدمة وتأثيرها على قوة الاختبار الإحصائي والدلالة العملية حيث اتفقت الأدبيات على التأثيرات السلبية للمعالجات بالطرق التقليدية على قوة الاختبار الاحصائي والتحيز والدلالة العملية.

- يؤمل أن تثير نتائج هذه الدراسة اهتمام المسئولين عن تطوير واستخدام المقاييس التربوية والنفسية والاستبانات لأغراض المسح مثل المسوح السكانية أو الدّراسات الحقلية مثل دراسات بمجال الصحة بإتّباع الطرق المناسبة لاستخدام الأسلوب الأمثل في معالجة مشكلة البيانات المفقودة دون حذف البيانات أو تغييرها، من خلال التحقق من النتائج المعمَمة، التي خرجت بها بعض الدّراسات السّابقة، حول أفضليّة طرق التعويض سواء المتعدد أو الانحدار أو تعظيم التوقعات. في الحصول على النتائج الإحصائية التي تتميز بقوة الاختبار الإحصائي لها بالظروف الاختبارية، من خلال دراسة التأثير لطرق التحليل وتفاعلها مع نسبة الفقد وحجم العينة والأسلوب الاحصائي للمعالجة.

- تأتي هذه الدراسة كاستجابة مُلحّة لكثرة الاعتماد على الطرق التي تعتمد على الحذف بالرغم من وجود اتجاه حديث وخاصّة من جمعية علم النفس الأمريكية بالاعتماد على الطرق الحديثة للتعامل مع البيانات المفقودة وأيضاً وجود اتجاهات حديثة مثل المعهد العربي للتدريب والبحوث الإحصائية (٢٠٢٣) باستخدام الطرق الحديثة الثلاث (التي حددتها الدراسة)، في بحوث ودراسات المسوح الميدانية، وبالتالي يتوقع من الدراسة الحالية توفير دليل مرجعي علمي لاستخدام الطرق الثلاث واستفادة المطوّرين والمحللين الاحصائيين من أسلوب معالجة البيانات بما يُسهم في

ISSN: 2537-0464 eISSN: 2537-0472

دقة الإجراءات العلمية في تحليل الدراسات والبحوث الكمية.

مصطلحات الدراسة:

١- البيانات (Data).

"معنى كلمة data هي جمع كلمة datum فإن قاموس أوكسفورد للغة الإنجليزية يقر بأن صيغة الجمع هذه وهي كلمة data تستعمل عادة بوصفها مفرداً"(العتيبي ٢٠١٢، ٥) وتعرف بأنها "هي مجموعة القيم التي يتم جمعها من مفردات المجتمع أو العينة الخاصة لخاصية (متغير) معينة" (خوارزم العلمية ٢٠١٢، ١٥). وعرفها العراقي بأنها "المادة الخام التي تحسب منها الإحصائيات وهي عبارة عن مجموعة من الحروف أو الكلمات أو الأرقام أو الرموز أو المؤشرات أو الإحصائيات الأولية بشأن موضوع محدد وليس بالضرورة أن تكون البيانات في صيغة رقمية دائماً، فهي في الدارج تأخذ صيغاً غير رقمية" (٢٠١٦، ٨١)

٢- البيانات المفقودة (Missing Data).

يمكن تعريف البيانات المفقودة بأنها

" تعني بعض الفقد لنوع من البيانات حول الظاهرة المراد دراستها، وهذا النوع من الفقد يعيق قدرتنا على فهم وتفسير تلك الظاهرة التي تدرسها"

(Mcknight, Sidani, & Figueredo, 2007,2).

كما يعرفها Mirzaei بأنها القيم المفقودة من متغير واحد أو أكثر في مجموعة البيانات، والتي تحدث نتيجة جمع البيانات بطريقة غير صحيحة أو أخطاء عند إدخال البيانات، مما يؤثر سلباً على الخطوات التالية في المشروع والذي يؤدي في النهاية إلى بناء استنتاجات غير صحيحة (Mirzaei, Carter, Patanwala & Schneider, 2021).

ويعرفها الباحث بأنها ترك المستجيبون، بعض فقرات الاختبار أو المفردات في المقياس دون استجابة لأسباب مختلفة، أو لأخطاء إدخال البيانات، مما تؤثر على دقة نتائج عمليات التحليل والتفسير للبيانات المدروسة، والتي بدورها تؤثر على اتخاذ القرارات المناسبة.

٣- طرق معالجة البيانات المفقودة (Imputation).

مجموعة من الطرق التي يمكن استخدمها لمعالجة البيانات المفقودة واستبدالها بقيم أخرى، وهذه الطرق هي: الحذف و تنقسم إلى طريقة الحذف الكلى للبيانات المفقودة، وطريقة الحذف الجزئي لبعض البيانات المرتبطة بتحليل معين، وطرق التعويض وتنقسم إلى طرق التعويض الفردي مثل: التعويض بقيم المتوسط، وطريقة تحليل الانحدار، وطريقة تعظيم التوقعات، وطريقة التعويض المتعدد (Baraldi & Enders, 2010,5-7). كما يعرفها Zhou "إحلال البيانات المفقودة

ISSN: 2537-0464 eISSN: 2537-0472

بتقديرات مناسبة أو وضع تقدير كمي مناسب بدلاً من البيانـات المفقودة" (Zhou,) 2001, 16

ويمكن تعريفها كما عرفها (Little & Rubin,2014) على النحو التالي:

أ- تعظيم التوقعات ((Expectation-Maximization Algorithm (EM) تعرف بأنها التعويض للبيانات المفقودة بالإعتماد على عمليات التقريب المتتابع ضمن خوارزمية مكونة من خطوتين هما خطوة التوقع، وخطوة التعظيم، اعتماداً على اقتران الأرجحية العظمى، بهدف الحصول على تقدير للقيم المفقودة.

ب- الانحدار ((Regression Imputation (Reg حيث هذه الطريقة يتم التعويض عن البيانات المفقودة للفقرة بعمل معادلة انحدار لكل فقرة فيها بيانات مفقودة.

ج- التعويض المتعدد ((Multiple-Imputation (MI تعرف بأنها طريقة تعتمد على ثلاثة أطوار (phases) لتقدير القيمة المفقودة، تبدأ بطور التعويض الذي يتم فيه تقدير القيم المفقودة في البيانات غير التّامة، ثم تليها خطوة التحليل بالاعتماد على أسلوب تحليل الانحدار لتحديد القيمة الاولية لكل قيمة مفقودة، وأخيراً خطوة الربط ويتم فيها إيجاد الوسط الحسابي لجميع التقديرات الأولية لكل قيمة مفقودة.

٤- العينة (Sample).

"تمثل العينة المجتمع الأصلي وتحقق أغراض البحث وتغني الباحث عن مشقات دراسة المجتمع الأصلي، وهكذا نعرف أن العينة هي جزء من مجتمع البحث الأصلي، يختارها الباحث بأساليب مختلفة، وتضم عددا من الأفراد من المجتمع الأصلي" (عبيدات، عبدالحق، وعدس ٢٠١٠، ٩٥)

ويعرفها الباحث بأنها تجمع أو اختيار بعض أفراد المجتمع إما بطريقة احتمالية أو غير احتمالية، لإيضاح التصورات أو الاستنتاجات عن المجتمع موضع الدراسة.

وفي الدراسة الحالية تم استخدام أسلوب المحاكاة **(Simulation).** وهي " أسلوب لدراسة وعمل أمثلة لنظم العالم الواقعي من خلال عمل التجارب بالأرقام العشوائية، وتستخدم في حل مشاكل رياضية وإحصائية معقدة يصعب حلها بالطرق التحليلية" (كاريان، دوديويز، ١٩٩٩، ٣٨٩)

ويعرف حجم العينة إجرائيا.

٥- قوة الاختبار الإحصائي (Power of Test).

عرّفها النجار (٢٠١٥، ١٣٨) بأنها "قدرة الاختبار على رفض الفرضية الصفرية عندما تكون في حقيقة الأمر خاطئة.**Power of Test=1-β** ، ويمكن تحديد حداً حرجا للرفض أو القبول للفرضية الصفرية مع الأخذ بالحسبان الفرضية البديلة".

كلما كبرت β يضعف الاختبار، وحتى تصغر β نحتاج لتكبير α.

ISSN: 2537-0464 eISSN: 2537-0472

ويعرف قوة الاختبار الإحصائي إجرائياً بأنه القيمة التي قام الباحث بحسابها لاختبار ت لعينة واحدة أو لعينتين مستقلتين أو تحليل التباين الأحادي في حالة تساوي حجوم العينات بدلالة مربع ايتا

ويمكن تعريفها بأنه احتمال الكشف عن الدلالة الإحصائية، بمعنى قدرة الاختبار الإحصائي على رفض الفرض الصفري (العدم) الخاطئ، والقبول بالفرضية البديلة الصحيحة، وبهذا المعني فقوة الاختبار الإحصائي تساوي ١- الخطأ من النوع الثاني؛ أو = ١-β.

٦- الدلالة العملية.

"الأساليب التي يتم من خلالها معرفة حجم الفرق أو حجم العلاقة بين متغيرين أو أكثر ويسمى أحيانا بحجم التأثير (significance effect size) أو قوة التأثير (strength effect) في الواقع العملي" (منصور، ١٩٩٧، ٥٧)، ويعرف الأثر (Effect) بأنه مقدار ثابت (يضاف إلى أو يطرح من) درجات المجموعة نتيجة لفعل مستوى معين من مستويات المتغير المستقل في حالة تحقق افتراضين الأول يتساوى أثر أحد مستويات المتغير المستقل بالنسبة لجميع أفراد المجموعة التي تتلقى المعالجة التجريبية والثاني أن يساوي مجموع الأثر لجميع المعالجات المختلفة صفراً (باهي ، ٢٠٠٦)، ويقاس حجم الأثر (Effect Size) بنسبة تباين المتغير التابع التي ترجع للمتغير المستقل، أي أنه يبين قوة العلاقة بين المتغير التابع والمتغير المستقل، فهو يعطي الدلالة العملية للفروق الإحصائية أو العلاقات بين المتغيرات وما إذا كانت تلك الدلالة العملية كبيرة تبرر الأخذ بنتائجها، وبذلك يتميز عن الدلالة الإحصائية التي تهتم باحتمالية رفض الفرضية الصفرية من الناحية الإحصائية النظرية فقط (Ellis,2010).

الإطار النظري والدراسات والبحوث السابقة

المحور الأول: البيانات المفقودة.

البيانات المفقودة تشير إلى الردود التي لم يتم الحصول عليها من المشاركين أو المستجيبين في استخدام أداة القياس مثل (الاختبار أو الاستبيان في الدراسات المسحية)، وذلك بسبب عدم القدرة على جمع البيانات من بعض أفراد الجمهور المستهدف أو العينة (الصرايرة، ٢٠١٨). كما يعرف ليثيل وروبن (Little, & Rubin, 2014) البيانات المفقودة بأنها هي البيانات غير التامة والتي تُصعب التحليل الإحصائي. وبصورة عامة تُعد مشكلة الاستجابات المفقودة للأفراد على عدد من فقرات الاختبار من القضايا البحثية التي يوليها اهتماماً الباحثين لتأثيرها المحتمل في دقة تقدير خصائص الاختبارات وكما تكتسب دراسة مشكلة البيانات المفقودة أهميتها عندما يتعلق الأمر بالجانب النفسي أو التربوي وخاصة عند الاعتماد على

ISSN: 2537-0464 eISSN: 2537-0472

تلك المعالم المقدرة في اتخاذ القرارات المتعلقة بقبول أو رفض الفرضية الصفرية وفق النتائج التي يسفر عنها تطبيق المقاييس التربوية والنفسية (اللصاصمة، ٢٠١٨).
ويُعد (Yates) أول من عالج البيانات المفقودة من خلال استخدامه لأسلوب الحذف والذي يقوم على حذف كل الحالات التي تتضمن قيماً مفقودة والاكتفاء على حجم العينة المتبقي بعد الحذف، ثم ظهر أسلوب معالجة بيانات القيم المفقودة من خلال الاعتماد على التحليلات الإحصائية المتعلقة بالتباين والتباين المشترك على يد العالم بارتليت (Bartlett) عام ١٩٣٩م (الحيالي، ٢٠١٣).

وقد بذل العلماء جهودًا متواصلة لتقديم مجموعة متنوعة من الأساليب لمعالجة مشكلة البيانات المفقودة. على الرغم من اختلاف هذه الطرق، فإن الدراسات حول البيانات المفقودة أدت إلى البحث عن أسباب الفقد. يشير الأدب العلمي إلى ثلاثة أنواع من آليات فقد القيم. النوع الأول يتعلق بالفقد العشوائي التام، حيث يكون سبب الفقد غير مرتبط بالقيمة المفقودة أو بقيم المتغيرات الأخرى. النوع الثاني، المعروف بالفقد العشوائي، يتميز بأن القيم المفقودة لها علاقة بقيم المتغيرات الأخرى ولكنها غير مرتبطة بالقيم المفقودة نفسها. أما النوع الثالث الفقد غير العشوائي، فينشأ عن القيمة المفقودة نفسها ولا يرتبط بالقيم الأخرى (Little & Rubin, 2014) .بذل الإحصائيون جهودًا كبيرة لتقديم العديد من الأساليب لتقدير القيم للبيانات المفقودة، سواء كانت تلك القيم قد فقدت بشكل عشوائي تام، أو بشكل عشوائي، أو بشكل غير عشوائي (اللصاصمة، ٢٠١٦).

وقلما نجد دراسة تتعامل مع العينات في التصاميم البحثية سواءً في الاختبارات أو الاستبانات إلّا وبها نسبة من البيانات المفقودة (Peugh & Enders, 2004)، ومن المتوقع أن يكون للبيانات المفقودة تأثير متغير على حجم البيانات الأصلية، وإهمال التعامل مع هذه البيانات قد يؤدي إلى تأثير سلبي على النتائج (Finch, 2008). وقد يتجاهل بعض الباحثين التعامل مع البيانات المفقودة بحكم عدم الوعي بحجم المشكلة أو الميل نحو السهولة لاستكمال البحث، ولعدم المعرفة بأساليب معالجة البيانات المفقودة، فمنها ما يحتاج إلى عمليات إحصائية وخبرة في البرمجيات التي تحتوي بعض المعالجات لفقد البيانات، وهذا يشجع الباحث على إهمال البيانات المفقودة أو حذفها باستخدام الطرق القديمة، والتي تشكّل نقصاً في المعلومات مما يؤثر على التفسير الصحيح لسلوك فقرات الاختبار أو أداء المستجيبين (Little &Rubin,2014).

وفيما يلي عرضاً للأسباب والأنماط ونبذة عن طرق وأساليب التعامل مع البيانات المفقودة، بالتركيز على الطرق الثلاث التي اعتمدت عليها الدراسة الحالية، حيث يشير الأدب إلى تنوع طرق تقدير القيم المفقودة في اساليب التقدير، فمنها ما يعتمد على إجراءات الحذف ومنها ما يعتمد على إجراءات القيمة التعويضية، وبالدراسة

ISSN: 2537-0464 eISSN: 2537-0472

الحالية اعتمدت على ثلاث طرق للتعويض هي (Expectation-Maximization (Algorithm (EM) طريقة لوغاريتم تعظيم التوقعات في تحديدها للقيمة التعويضية للبيانات المفقودة على عمليات التقريب المتتابع ضمن خوارزمية مكونة من خطوتين هما خطوة التوقع، وخطوة التعظيم، اعتماداً على اقتران الأرجحية العظمى، بهدف الحصول على تقدير للقيم المفقودة. في حين تتضمن طريقة التعويض المتعدد ((Multiple-Imputation (MI) ثلاثة أطوار (phases) لتقدير القيمة المفقودة، تبدأ بطور التعويض الذي يتم فيه تقدير القيم المفقودة في البيانات غير التّامة، ثم تليها خطوة التحليل بالاعتماد على أسلوب تحليل الانحدار لتحديد القيمة الاولية لكل قيمة مفقودة، وأخيراً خطوة الربط ويتم فيها إيجاد الوسط الحسابي لجميع التقديرات الأولية لكل قيمة مفقودة. في حين التعويض من خلال الانحدار ((Regression Imputation (Reg حيث هذه الطريقة يتم التعويض عن البيانات المفقودة للفقرة بعمل معادلة انحدار لكل فقرة فيها بيانات مفقودة. (Little & Rubin,2014).

أسباب حدوث البيانات المفقودة.

من الضروري جداً فهم الأسباب التي تؤدي إلى ظهور البيانات المفقودة، حيث يرتبط ذلك بالقرارات المتعلقة بالأنماط العامة لآليات الفقد والطرق المستخدمة للتعامل معها. حيث أشار (Mislevy & Wu , 1988) إلى أن الاستجابات المفقودة في أي اختبار أو أي أداة قياس، قد تكون نتيجة لتصميم الاختبار أو نتيجة لاختيار المفحوص نفسه. ومن جهة أخرى، يعتقد لودلو وأولري (Ludlow & O,leary , 1999) أن البيانات المفقودة قد تحدث بسبب عدم الإجابة، سواء كان ذلك بسبب حذف الاستجابة أو عدم القدرة على الوصول إليها. الاستجابات المحذوفة هي تلك التي تم تجاوزها بطريق الخطأ من قبل المفحوصين، أو فشل في الإجابة عليها، أو لم يتمكن المفحوص من الوصول إليها ولم يقض وقتاً كافياً في أداء الاختبار بسبب نقص الوقت أو الاهتمام. بشكل عام، تتأثر البيانات التي يتم جمعها في البحوث والدراسات النفسية والتربوية بشكل كبير بسبب عدم استجابة بعض المفحوصين لبعض فقرات أداة القياس، بغض النظر عن سبب عدم الاستجابة، وهذا يؤدي إلى وجود بيانات مفقودة، وبالتالي الحصول على بيانات ناقصة يؤثر على فعالية التفسيرات والتخاذ القرارات المناسبة.(Little & Rubin,2014) وعند الرجوع إلى الأدب السيكو متري، يمكن التفريق بين نوعين من عدم الاستجابة وهما:

- عدم الاستجابة الكاملة، حيث لا يقدم الفرد أي استجابة لأي من فقرات الاختبار. هذا يحدث عندما يكون الفرد غير متواجد أو يرفض المشاركة بسبب القلق أو عدم الأمان حول سرية المعلومات، أو عدم الحضور بسبب العمر أو الحالة الصحية.

ISSN: 2537-0464 eISSN: 2537-0472

• عدم الاستجابة للفقرة، حيث يستجيب المفحوص لبعض الفقرات ويترك بعضها دون إجابة. هنا يتوفر بعض البيانات المستجابة وبعض البيانات المفقودة. بمعنى أن المفحوص يشارك ولكنه لا يستجيب لبعض الفقرات. هذا يشمل تخطي الفقرات، أي أن المفحوص تركها فارغة، إما لأنها تتطلب جهدًا أكبر من اللازم للإجابة عليها، أو لأنه لا يعرف الإجابة لسبب ما، أو لأنه قد تخطاها بغير قصد (سقطت سهواً) أو بقصد لكونها محرجة له، أو لأن الوقت لم يكن كافياً للإجابة عليها فتركت فارغة.

ومن خلال ما سبق، يلخص الباحث أسباب حدوث الفقد للبيانات برفض المفحوصين الإجابة على بعض الفقرات لأنها تتعلق بقضايا خاصة بهم أو يعتقد بأنها تسبب تهديدا لهم، أو لأنها تتناول موضوعًا حساسًا، وقد لا تكون مناسبة للمفحوصين من حيث المحتوى أو لا تنطبق عليهم، أو بسبب طول المقياس الذي يسبب الإرهاق للمفحوصين أو الملل، أو ضيق الوقت، أو الفقد أثناء عملية الإدخال للبيانات بعد جمعها أو اثناء ترميزها، أو بسبب الحروب والكوارث، أو ارتفاع التكلفة، أو الإهمال أو لعدة عوامل متداخلة (العتيبي، ٢٠١٨).

ويرى الباحث أنه من الضروري معرفة سبب الفقد عن طريق الأنماط وآلية لفقد، لاختيار الطريقة المناسب لمعالجة تلك البيانات المفقودة.

آلية فقد البيانات.

حسب الاتفاق بين الباحثين مثل (Little & Rubin, 2014; Enders, 2010; Graham, 2012) هناك ثلاثة تصنيفات لآليات فقد البيانات وهي:

١- الفقد عشوائياً ((Missing at Random (MAR).

تعني أن الفقد فيها مرتبط بطريقة مباشرة أو غير مباشرة بمتغيرات خارجية، فإذا كانت هناك فقرة مفقودة وهذه الفقرة تقيس السمة الكامنة على مستوى المقياس ككل كبقية الفقرات الأخرى، فالسبب في الفقد يعزى إلى بعض المتغيرات الخارجية، ويقال إن البيانات فقدت بشكل عشوائي إذا كانت البيانات المفقودة لا تعتمد على قيم المتغير المستقل نفسه الذي يحتوي على هذه القيم المفقودة. ولكنها تعتمد على متغير (أو متغيرات) مستقلة أخرى.

٢- الفقد العشوائي بالكامل (Missing Completely at RandoM)) (MCAR).

تعني أن الفقرة مفقودة بسبب العشوائية المحضة فقط، لا لأي سبب آخر، أي أن آلية الفقد لا ترتبط بمتغيرات الدراسة ويقال إن البيانات فقدت بشكل عشوائي تام إذا كانت البيانات المفقودة لا تعثمد على قيم المتغير المستقل نفسها الذي يحتوي على هذه القيم المفقودة أو أي متغير مستقل آخر في قاعدة البيانات. وهذا الفقد ما تم اعتماده من قبل الباحث في هذه الدراسة.

٣- الفقد غير العشوائي ((Missing Not at Random (MNAR)**.**

ISSN: 2537-0464 eISSN: 2537-0472

تعني أن البيانات المفقودة مرتبطة بالمتغير نفسه، فلها علاقة إما بمعالم الفقرة (صعوبة، تمييز، تخمين) أو بقدرة الأفراد ويقال إن البيانات فقدت بشكل غير عشوائي إذا كانت البيانات المفقودة في متغير مستقل معين تعتمد على قيم المتغير نفسها.

طرق التعامل مع البيانات المفقودة.

خلال العقدين الأخيرين، ازداد البحث في مجال التعامل مع البيانات المفقودة، وشهدنا تحسنًا واضحًا في الأساليب والإجراءات الإحصائية التي تتعامل مع البيانات المفقودة. هناك العديد من الطرق المتنوعة لمعالجة البيانات المفقودة، والتي تتم قبل بدء عملية التحليل. هذه الطرق تنقسم إلى نوعين رئيسيين: الحذف وحساب القيم البديلة، فيما يلي توضيح للطرق المستخدمة في كل منهما.

أولا: الطرق القائمة على الحذف.

الباحثون يستخدمون هذه الأساليب للتعامل مع البيانات المفقودة بهدف تقديمها كبيانات كاملة، والتي غالبًا ما تؤدي إلى نتائج متحيزة وغير فعالة بالدرجة المأمولة التي تعطي نتائج يمكن الوثوق بها

(Little & Rubin, 2014) تتضمن هذه الطرق ما يلي:

١- **حذف الحالة** (Case Deletion)**.**

وتصنف إلى طريقتين أو أسلوبين من حيث طريقة التحليل، وهما:

أ- تحليل الحالة الكاملة (Complete-Case Analysis)

ب- وتحليل الحالة المتوفرة (Available-Case Analysis)

وهاتان الطريقتان الأكثر استخداماً لمعالجة البيانات المفقودة لسهولة التعامل مع هذه البيانات، وعملهما يتلخص في التخلص من المفحوصين الذين لديهم بيانات مفقودة، وهما موجودان كخيارين من خيارات أخرى في برنامج (SPSS) (Little & Rubin, 2014). ويتم التمييز بين هاتين الطريقتين اعتماداً على نوع التحليل، فإذا كان المطلوب إيجاد تحليل الانحدار المتعدد أو إيجاد تحليل التباين، فإن الطريقة المستخدمة هي: تحليل الحالة الكاملة Complete-Case Analysis، أمّا إذا كان المطلوب إيجاد الارتباط بين متغيرين فإن الطريقة المستخدمة هي: تحليل الحالة المتوفرة Available-Case Analysis (Graham, 2009). ويرى شيفر وجراهام (Schafer & Graham, 2002) بأن هاتين الطريقتين من أقل الطرق كفاءةً للتعامل مع البيانات المفقودة.

أ- **تحليل الحالة الكاملة** ((Complete-Case Analysis (CC)

هذه الطريقة موجودة كخيار منتقى من خيارات أخرى في برنامج (SPSS)، ونحصر الاهتمام هنا بالمفحوصين الذين لهم استجابات لكل الفقرات، أي يتم التخلص

ISSN: 2537-0464 eISSN: 2537-0472

من أي مفحوص له قيمة مفقودة؛ لأي من الفقرات، وتحليل المفحوصين الذين استجابوا لجميع الفقرات، وبالرغم من أن كل فقرة مفقودة تكون عليها نسبة مئوية صغيرة من الاستجابات، فالجزء الكبير من المفحوصين لا يتم حذفه، فهذا التخفيض في حجم العينة يؤثر سلباً في القوة الإحصائية، وقد يجعل مسألة تمثيل العينة للمجتمع في موضع شك، ولكي تكون العينة الباقية على الأرجح ممثلة، فإنه ينبغي أن يكون عدد المفحوصين الذين تم حذفهم قبل التحليل قليل، ومع هذا فإن ذلك يؤدي إلى خسارة للمعلومات بسبب حذف بعض المفحوصين، وبالتالي يزداد احتمال التحيز (Pigott, 2001)، وهذه الطريقة حرجة جداً في العينات الصغيرة، وتسمى هذه الطريقة بتحليل الحالة الكاملة Complete-Case Analysis لأنها تحلل المفحوصين الذين لديهم استجابات كاملة على جميع الفقرات، وتسمى أيضاً List-wise deletion لأنها تزيل وتتخلص من المفحوصين الذين لديهم استجابات ناقصة على الفقرات (Witta, 2000).

ب- **تحليل الحالة المتوفرة:**((Available-case analysis (AC)

هذه الطريقة موجودة كخيار منتقى من خيارات أخرى في برنامج (SPSS)، وتهتم هذه الطريقة بالمفحوصين الذين لهم استجابات على كلا المتغيرين إي أن البيانات متوفرة على أزواج المتغيرات، فإذا أردنا حساب معامل الارتباط بين (10) عشرة متغيرات لإيجاد مصفوفة التباين المشترك (covariance matrix)، فإن عدد معاملات الارتباط بين الفقرات تساوي (45) معامل ارتباط، أي أن عناصر المصفوفة تساوي (45) عنصر، وكل عنصر يشكل معامل ارتباط بين متغيرين، وقيمته تأتي من المفحوصين ، ولكن نفس المفحوصين قد لا ينتمون إلى كل ارتباط في المصفوفة لأن نمط البيانات المستجابة والمفقودة لكل مفحوص قد تختلف، فعلى سبيل المثال مصفوفة الارتباط بين بعض المتغيرات قد تتضمن عدد كبير من المفحوصين، ولكن على المتغيرات الأخرى قد تتضمن مصفوفة الارتباط عدد صغير من المفحوصين، مما يجعل المجموعات غير متساوية فيصعب حساب درجات الحرية وإذا تم اختيار حجم العينة الأقل فهذا يؤثر على القوة الإحصائية (Acock, 2005)، والتحليلات الإحصائية بهذه الطريقة تصبح غير مستقرة، وتنتج قيم غير ملائمة خاصة عندما يكونان المتغيران مرتبطين معاً فإن معلومات أحدهما يحسن تقديرات المتغير الآخر، وهذا يعطي تقديرات متحيزة (Witta, 2000). وتسمى هذه الطريقة بِـ (pair-wise deletion) لأنه يتم التخلص من المفحوصين الذين لديهم استجابات مفقودة على كلا المتغيرين، وتسمى هذه الطريقة أيضاً بِـ (pair- wise inclusion) . لأنه يتم إدراج المفحوصين الذين لديهم الاستجابات على كلا المتغيرين.

ISSN: 2537-0464 eISSN: 2537-0472

٢- إعادة الوزن:

يمكن استخدام طريقة "إعادة الوزن" لتقليل التحيز، بحيث يتم حذف البيانات المفقودة، وإعطاء أوزان للبيانات المستجابة سواءً كانت خاطئة أو صحيحة، ويتم إيجاد قيم لهذه الأوزان من خلال النماذج الرياضية، أو من خلال الاحتمالات (بني عواد والمومني، ٢٠١٨).

٣- حذف الصفوف والأعمدة:

تعتبر هذه الطريقة الأكثر شيوعاً للتعامل مع القيم المفقودة. يتم حذف الصف أو العمود في حال كان يحتوي على سبيل المثال نسبة ٧٠-٧٥٪ من القيم المفقودة. ينصح بهذه الطريقة فقط عند وجود عدد ضخم من البيانات، فهي تعتبر خيار غير جيد إذا كانت النسبة المئوية للقيم المفقودة عالية، على سبيل المثال ٣٠٪ مقارنةً بمجموعة البيانات بأكملها (Kaiser, 2014).

وبالدراسة الحالية الحالية تُعد جميع الطرق السابقة تقليدية وتبين قلة كفاءتها مقارنة بالطرق الحديثة التي صُنّفت ضمن الجزء الثاني القائم على التعويض، ولكن ليس كل طرق حساب القيم التعويضية طرقاً حديثة.

ثانياً: الطرق القائمة على احتساب قيمة تعويضية.

وتعتمد هذه الطرق على التقدير لتعويض البيانات المفقودة وهي:

١- **حساب قيمة تعويضية واحدة:** (Single Imputation)

عندما يتوفر بيانات جزئية يتم استبدال القيم المفقودة بقيم معقولة من خلال البيانات الجزئية ومواصلة التحليل، فالتعويض قائم على عدم إهدار أي معلومات من أفراد العينة، والاحتفاظ بكل العينة بمعنى عدم خسارة أي معلومات ناجمة عن النقصان في حجم العينة، وإذا احتوت البيانات المتوفرة بيانات مفيدة للتنبؤ بالقيم المفقودة، بالإمكان إجراء التعويض واستغلال هذه البيانات، والمحافظة على دقة عالية من النتائج، فالتعويض ينتج مجموعة بيانات كاملة يمكن تحليلها بالطرق والبرامج الحاسوبية (Schafer & Graham, 2002) .

٢- **حساب قيمة تعويضية من خلال الانحدار:**(Regression Imputation)

في هذه الطريقة يتم تعويض البيانات المفقودة للفقرة بعمل معادلة انحدار لكل فقرة فيها بيانات مفقودة، والعلامات المتنبئ بها من معادلة خط الانحدار تستخدم لاستبدال البيانات المفقودة (Kalaycioglu, Copas, King, Omar, 2016). كما تتضمن طريقة الانحدار التوصل إلى معادلة انحدار بطريقة المربعات الصغرى، ويكون فيها المتغير ذو البيانات المفقودة متغيراً تابعاً وتستخدم فيها المتغيرات الأخرى ذات العلاقة والتي تكون قوة ارتباطها بالمتغير متوسطة على الأقل بوصفها متغيرات متنبئة بالقيم المفقودة للمتغير (فهمي، ٢٠٠٩). وتتباين أنواع الانحدار

ISSN: 2537-0464 eISSN: 2537-0472

حسب طبيعة المتغير ما يؤدي إلى تباين طريقة تقدير القيم المفقودة من متغير إلى آخر في نفس الدراسة (Musil, Warner, Yobas, Jones,2022). ولهذه الطريقة العديد من العيوب منها أنها قد تؤدي إلى تعديل في البيانات أكثر من اللازم، كذلك تفترض وجود علاقة ارتباط مع المتغير الذي يحتوي على بيانات مفقودة. ولذلك إذا كان الارتباط ضعيفا فإنه لا يمكن التنبؤ بالبيانات المفقودة. وعلى الرغم من عيوبها إلا أنها تصلح في الحالات التي تنتشر (العتيبي، ٢٠١٨).

المحور الثاني: قوة الاختبار الإحصائي والدلالة العملية.

قوة الاختبار الإحصائي (Power of test).

تعتبر قوة الاختبار (Power of test) الإحصائي مقياساً لقدرة الاختبار على اكتشاف الفروق الحقيقية بين المجموعات المختلفة في الدراسة، حيث تُعرّف بأنها " قدرة الاختبار الإحصائي على رفض الفرض الصفري الخاطئ، وبهذا المعني فقوة الاختبار الإحصائي تساوي (١- الخطأ من النوع الثاني، **(1-β)**) (علام، ٢٠١٠).

ولذا ففي حالة قبول الفرض الصفري فهناك احتمال أن يكون هناك خطأ من النوع الثاني، وتظهر أهمية الكشف عن قوة الاختبار الإحصائي للتأكد من تقليل احتمال رفض الفرض الصفري الخاطئ ، وهناك اتفاق بين المهتمين على أن قوة الاختبارات الإحصائية المقبولة يجب ألا تقل عن ٠.٨٠ بمعنى أن نسبة الخطأ من النوع الثاني يجب ألا تزيد عن ٠.٢٠ ففي هذه الحالة وإذا فرضنا أن مستوى الدلالة سيكون ٠.٠٥ فإن نسبة الخطأ من النوع الثاني إلى الخطأ من النوع الأول تعني أن احتمالية الخطأ من النوع الثاني ستكون أربعة أمثال نسبة الخطأ من النوع الأول وهي نسبة كبيرة، لكن الخطأ من النوع الثاني يعتبر أقل خطورة من الخطأ من النوع الأول، ويتشدد البعض في تقرير قوة الاختبارات الإحصائية ويعتمد على أن نسبة الخطأ من النوع الثاني يجب ألا تتعدي ٠.٠٥ بمعنى أن قوة الاختبار يجب ألا تقل عن ٠.٩٥(فهمي، ٢٠٠٩).

"ومن معايير اختيار الطرق الإحصائية الملائمة لتحليل البيانات، قوة الاختبار حيث يعتبر هذا المعيار محكا تفضيلياً في حالة استخدام أكثر من أسلوب إحصائي لترجيح كفة الأسلوب أو الاختبار الأقوى لرفض الفرضية الصفرية عندما تكون خاطئة" (طه، وفلمبان، ٢٠٢٤، ٤٢).

فوائد حساب حجم التأثير:

١- حساب حجم التأثير يسهل من التحليل المتلاحق حتى مع دمج تقارير مختلفة.

٢- ملاحظة حجم التأثير في المجالات الأدبية التي بها يتم عن طريق الأبحاث المتتابعة استنباط توقعات محددة، وذلك عن طريق التكامل بين النتائج الملاحظة مع الدراسات السابقة.

ISSN: 2537-0464 eISSN: 2537-0472

٣- ربما هو الأهم، تفسير حجم التأثير في دراسة ما يسهل من أنه يمكن تقييم أن البحث تحت الدراسة يتبع قواعد أدبية محددة. (Gordon Howard R. D. , : p7 2001).

وبالدراسة الحالية تم الاعتماد على حساب مربع إيتا (حجم التأثير) لقياس قوة الاختبار الاحصائي حيث أنه بحسب ما ذكره. (Huston , 1993; Fan , 1999; Onwuegbuzie et al ,2003) يمكن استخدامه في تحليل القوة الإحصائية؛ لتحديد عدد العناصر المطلوبة في دراسة معينة، كما يؤكد (Steyn & Ellis, 2009) على أنه يمكن استخدامه لحساب قوة الاختبار الاحصائي ولتحديد قوة الاختبار لحجم العينة المخطط لها أو للحصول على حجم العينة للقوة المعطاة. ويمكن تقسيم مقاييس الدلالة العملية (حجم التأثير) إلى قسمين هما : مقاييس حجم التأثير المستخدمة مع اختلافات المتوسطات بين المجموعات وتدل على حجم الأثر للمتغير المستقل على المتغير التابع كما في اختبار (t) للعينتين واختبار (F) لأكثر من مجموعتين، ومقاييس حجم التأثير المستخدمة لقياس العلاقات وقوة العلاقة بين متغيرات الدراسة وتعتمد حساب التباين مثل R^2 ومربع إيتا Eta Squared أو مربع أوميجا Omega Squared .

الدراسات والبحوث السابقة:

في دراسة (Fu, et al, 2021) هدفت إلى المقارنة بين كفاءة مجموعة من طرق وأساليب معالجة البيانات المفقودة بقاعدة البيانات الدولية للتربة (UNSODA)، ولمعالجة بيانات خصائص التربة المفقودة وتم استخدام ثلاث طرق قائمة على التعلم الآلي (انحدار الغابات العشوائية (RF)، وانحدار ناقل الدعم (SVR)، وانحدار الشبكة العصبية الاصطناعية (ANN)، وطريقتان قائمتان على الإحصائيات، أي المتوسط والتعويض المتعدد (MI) ، تم استخدامها لحساب بيانات خصائص التربة المفقودة، وتم إجراء محاكاة للقيمة المفقودة وتقييمها كمياً. بعد ذلك، تم إجراء الاختبارات اللامعلمية والانحدار الخطي المتعدد لإجراء تقييم نوعي لموثوقية طرق التضمين الخمسة هذه. أظهرت النتائج أن (RMSEs) و(MAEs) لجميع الميزات تقلبت ضمن النطاقات المقبولة. وتم حساب خصائص التربة باستخدام الأساليب الاحصائية (MI) أدنى (RMSEs) و(MAEs)؛ كلتا الطريقتين جيدتان في شرح تباين البيانات. حيث انخفض الخطأ المعياري ومعامل التباين والانحراف المعياري بشكل ملحوظ بعد التعويض، ولم تكن هناك فروق ذات دلالة إحصائية قبل وبعد التضمين وتشير هذه الدراسة إلى أن أساليب (RF) و(MI) قد تكون أفضل لإدراج البيانات المفقودة.

ودراسة العتيبي (٢٠٢٢) هدفت إلى تحديد أكفأ طرق التعويض عن البيانات

ISSN: 2537-0464 eISSN: 2537-0472

المفقودة، وذلك من خلال مقارنة طريقة التعويض بأخر قيمة ملاحظة، وطريقة هوت دك لتقدير البيانات المفقودة؛ ولتحقيق هدف الدراسة تم استخدام المحاكاة لتوليد أرقام عشوائية بأحجام عينات مختلفة (٣٠، ٦٠، ٩٠)، حيث تم بناء نموذج محاكاة لأسلوب تحليل التباين الأحادي (في حالة تساوي أحجام العينات بين المجموعات، وفي حالة عدم تساوي أحجام العينات بين المجموعات) وتتوافر فيه الافتراضات الخاصة بأسلوب تحليل التباين الأحادي. وتم حساب القوة الإحصائية للنموذج التام للحالات المذكورة. ثم تم عمل فقد لهذه البيانات بنسب فقد بلغت (١٥%)، و (٢٥%)، و (٤٠%)، على التوالي وفق آلية الفقد (MCAR)، وتحت نمط فقد (النمط العام). وبعدد ذلك تم تقدير القيم المفقودة باستخدام طريقتي التعويض. ثم إجراء اختبار تحليل التباين الأحادي للنماذج المقدرة، وحساب القوة الإحصائية لتلك النماذج للحالات المذكورة، ثم تتم المفاضلة بين طريقتي التعويض بآخر قيمة ملاحظة وهوت دوك لتقدير البيانات المفقودة، وذلك من خلال مقارنة القوة الإحصائية لتلك النماذج المقدرة بقوة النموذج الإحصائي التام. وأظهرت نتائج الدراسة أن طريقة التعويض بآخر قيمة ملاحظة (LOCF) أكفا من طريقة هوت دك (HOT DECK) بجميع حالات نسب الفقد وأحجام العينات المختلفة، إذ تعطي طريقة (LOCF) قيما تقديرية تؤدي إلى نتائج أقرب إلى النموذج التام وفقا لمعيار قوة الاختبار الإحصائي للنموذج.

ودراسة (Adjei, Rose, 2022) هدفت إلى تقديم نظرة ثاقبة للبيانات المفقودة عبر مجالات الدراسة المختلفة وتتناول بعض التحديات الإحصائية (قوة الاختبار الاحصائي) للبيانات المفقودة من خلال دراسات المحاكاة والتطبيق على مجموعات البيانات الحقيقية. تم محاكاة لظاهرة التسرب في تسلسل الحمض النووي الريبي النووي (scRNA) من خلال تحليل مقارن لبعض تقنيات تسلسل الحمض النووي الريبوزي (scRNA) الموجودة، باستخدام دراسات المحاكاة لتقييم ما إذا كان من المناسب معالجة مسألة البيانات المفقودة باستخدام طريقة التعويض التقليدية، أو التضمين والتوقع، أو الأسلوب القائم على عدم الافتراض بالعشوائية. واقترحت استراتيجية فعالة لمعالجة مسألة عدم التوازن في البيانات (سواء كانت معتدلة أو غير متوازنة للغاية) من خلال الجمع بين أخذ العينات العشوائية مع استراتيجيات الترجيح المختلفة. ونستنتج بشكل عام، بناءً على النتائج التي توصلت إليها أن الغياب لا يعني دائماً نقصاً في المعلومات ولكنه مثير للاهتمام يحتاج إلى التحقيق.

ودراسة جوريتزكو (Goretzko, 2022) هدفت إلى تقييم طرق معالجة البيانات المفقودة مثل (خوارزمية تعظيم التوقعات (an Expectation-Maximization Algorithm) ومطابقة المتوسط التنبؤي، والتعويض المتعدد باستخدام الغابات العشوائية ضمن التضمينات المتعددة بواسطة إطار المعادلات المتسلسلة (MICE) بالإضافة إلى الحذف الزوجي، على قوة ودقة محكات

ISSN: 2537-0464 eISSN: 2537-0472

وقرارات الاحتفاظ بالعامل في تحليل العوامل الاستكشافية تمت محاكاة البيانات لأحجام عينات مختلفة، وأعداد العوامل، وأعداد المتغيرات الواضحة (المؤشرات)، وآليات البيانات المفقودة ونسب القيم المفقودة. وتبين أنه لم يكن لآلية البيانات المفقودة تأثير يذكر على الدقة وكان أداء الحذف الزوجي جيداً نسبياً في بعض الحالات، خاصة حالات العينات الصغيرة وعندما تم استخدام بيانات المقارنة لتحديد عدد العوامل، كان التضمين العشوائي (للغابات العشوائية) (Random forest Imputation) أفضل من طرق البيانات المفقودة الأخرى.

أوجه الاستفادة من الدراسات والبحوث السابقة.

استفاد الباحث من الدراسات والبحوث السابقة في بلورة الإطار النظري والربط والتحليل مع نتائج الدراسة الحالية، وتشكيل وتدعيم مشكلة الدراسة بناء على نتائج وتوصيات الدراسات السابقة، والاستفادة من الاساليب التي استخدمتها الدراسات السابقة، وتحديد طرق معالجة البيانات المفقودة ومقارنة نتائج الدراسة الحالية مع نتائج الدراسات السابقة، والتعرف على طريقة الكشف عن أثر طريقة المعالجة ونسبة الفقد وحجم العينة على قوة الاختبار الاحصائي والدلالة العملية. وتتميز الدراسة بالأصالة كونها من الدراسات القليلة – في حدود علم الباحث – التي ربطت بين معالجة القيم المفقودة بالطرق الثلاث معاً ونسب الفقد وحجم العينة لتأثيرها على قوة الاختبار الإحصائي والدلالة العملية لاختبار (ت، ف) لمجموعة واحدة ومجموعتين وثلاث مجموعات، كموضوع انفردت به الدراسة الحالية عن غيرها من الدراسات السابقة في البيانات المفقودة والتي اطلع عليها الباحث.

منهج وإجراءات الدراسة

منهج الدِّراسَة

هدفت الدِّراسَة إلى استقصاء تأثير طرق معالجة البيانات المفقودة ونسبة الفقد وحجم العينة على قوة الاختبار الإحصائي والدلالة العملية، ولتحقيق ذلك تمّ اتباع المنهج التجريبي، القائم على توليد البيانات الافتراضية بأسلوب المحاكاة للواقع لما توفره تلك البيانات من ظروف معيارية يصعب الحصول عليها في حال استخدام البيانات الواقعية.

بيانات الدِّراسَة:

تم توليد بيانات الدراسة باستخدام برنامج (WINGEN3) بهدف توليد بيانات الدراسة؛ حيث أعتمد على التوزيع الطبيعي للقدرة بوسط حسابي مقداره صفر وانحراف معياري مقداره واحد، في صورة الاختبار يتكوّن من (٤٠) فقرة للمستوى المنخفض من أجل أن يُراعي التمايز بين (مجموعتين أو أكثر)، في حين اعتمد التوزيع الطبيعي للقدرة بوسط حسابي مقداره نصف وانحراف معياري مقداره واحد

ISSN: 2537-0464 eISSN: 2537-0472

في صورة الاختبار للمستوى المرتفع؛ لعشرة مستنسخات(Replications) بحجم (٢٠٠، ٦٠٠، ١٠٠٠) لكل مستنسخة. كما اعتمد التوزيع الطبيعي لمعلمة الصعوبة بوسط حسابي مقداره صفر وانحراف معياري مقداره واحد، واعتمد التوزيع الطبيعي اللوغاريتمي لمعلمة التمييز بوسط حسابي مقداره (٠.٦٠) وتباين مقداره (٠.٢٥) في صورة الاختبار للمستوى المنخفض، في حين جرى اعتماد التوزيع الطبيعي لمعلمة الصعوبة بوسط حسابي مقداره نصف وانحراف معياري مقداره واحد، واعتماد التوزيع الطبيعي اللوغاريتمي لمعلمة التمييز بوسط حسابي مقداره (٠.٨٠) وتباين مقداره (٠.٢٥) في صورة الاختبار للمستوى المرتفع، كما تم تدريج الدرجة الكلية للاختبار بحيث يظهر التمايز بين مجموعات الدّراسة، وتم إنشاء (٣) ملفات لكل حجم عينة على حِدة وتُمثل نتائج تحليل درجاتها وفق الأسلوب الإحصائي المستخدم درجات ما قبل الفقد.

اجراءات تطبيق الدّراسة (المحاكاة).

تم إجراء دراسة المحاكاة وفق الخطوات التسلسلية التالية:

١- توليد ثلاثة بيانات (رقمية) بالطريقة العشوائية تتبع التوزيع الطبيعي القياسي وبحجم العينة (٢٠٠، ٦٠٠، ١٠٠٠) (إنشاء ثلاثة ملفات) متضمّنة مجموعة واحدة ومجموعتين وثلاث مجموعات، بحيث توفر جميع الافتراضات اللازمة للأساليب الإحصائية (اعتدالية التوزيع الطبيعي وتجانس التباين وتم حساب قوة الاختبار الإحصائي بدلالة مربع ايتا والكشف عن دلالة الفروق وفقاً للمجموعات).

٢- تم عمل فقد للبيانات بالملفات الثلاث بنسب الفقد المختلفة ١٠%، ٢٠%، و٤٠% وبالتالي إنشاء (٩) ملفات باستخدام برنامج (R) والتأكد من تحقق شروط آلية الفقد التام (MCAR) من حيث توزيع البيانات المفقودة (١٠%، ٢٠%، ٤٠%) وحجم العينة (٢٠٠، ٦٠٠، ١٠٠٠) كما تتضح النتائج بجدول (١).

جدول (١) ملخصاً لنتائج اختبار Little's MCAR Test لتقدير العشوائية التامة للبيانات المفقودة على عينات دراسة المحاكاة

نسبة الفقد	طريقة المعالجة	٢٠٠			٦٠٠			١٠٠٠		
		$\chi 2$	DF	الدلالة	$\chi 2$	DF	الدلالة	$\chi 2$	DF	الدلالة
%١٠	EM	٦٩٣٤.٤	٦٩٢٤	غير دالة	١١٥٧.٣	١١٥٧	غير دالة	١٧٧٣.٥	١٩٦٠	غير دالة
	Reg	٦٩٢١.٥	٦٨٦١	غير دالة	١٢٣١.٨	١٢٢٧	غير دالة	١٧٦٩.٥	١٩١٨	غير دالة
	MI	٦٩١١.٣	٧٤٣١	غير دالة	١٢١٩.٩	١١٨٥	غير دالة	١٨٠٩.٢	٢٠١١	غير دالة
%٢٠	EM	٦٣٤٩.١	٦٣٦٠	غير دالة	١١٥٨٢	١٢٤٧	غير دالة	١٧٢٠.٠٩	١٩٦٢	غير دالة
	Reg	٦٢٥٠.٧	٦٧٧٨	غير دالة	١٢٧٤.٣	١١٧٩	غير دالة	١٧٦٩.٥	١٥٤٨	غير دالة
	MI	٧٤٧٢.٣	٧٤٠٨	غير	١٢٥٨.٢	١١٨٦	غير	١٧١٦.٨	١٩٥٨	غير

ISSN: 2537-0464 eISSN: 2537-0472

					دالة				دالة
%٤٠	EM	٤٦٦٩.٦٥	٤٧٦٠	غير دالة	١٥٤٢.٣	١٢٧٤	غير دالة	١٧٦٢.٣	١٦٥٩	غير دالة
	Reg	٤٥٥٥.٢	٤٦٢٥	غير دالة	١٢١٨.٢	١١٩٠	غير دالة	١٨٢٤.٩	١٩٤٨	غير دالة
	MI	٤٨٦٢.٥	٤٦٠١	غير دالة	١١٧٢.٣	١١١٠	غير دالة	١٨٢٥.٧	١٩٨٨	غير دالة

يتضح من جدول (١) أن جميع قيم كاي تربيع غير دالة إحصائياً مما يدل على أن البيانات المفقودة في جميع حالات النسب المفقودة (١٠%، ٢٠%، ٤٠%) ولجميع احجام العينات (٢٠٠، ٦٠٠، ١٠٠٠) ولجميع المعالجات المستخدمة تتوزع بشكل عشوائي تام.

٣- تم استخدام ثلاث طرق حديثة لمعالجة البيانات المفقودة (تعظيم التوقعات، القيم التعويضية المتعددة، طريقة الانحدار) للتعويض عن القيم المفقودة وإنشاء ملفات ثلاث بعد الفقد لكل نسبة فقد ولكل حجم عينة وبالتالي كان مجموع الملفات بعد الفقد (٩) ملفات وبالتالي تم الانتهاء من تقدير القيم المفقودة باستخدام الطرق الثلاث.

٤- إجراء التحليلات الإحصائية والكشف عن الدلالة الإحصائية وحساب حجم التأثير ومربع ايتا للوصول إلى قيمة قوة الاختبار الاحصائي والدلالة العملية، وذلك للمفاضلة بين الطرق الثلاثة من حيث اقترابها للقيمة قبل الفقد.

٥- تكرار الخطوات السابقة ١٠٠٠ مرة لكل سيناريو.

٦- الاعتماد على المتوسط الحسابي بمخرجات البرنامج (١٠٠٠ مرة) في تقييم الطرق التعويضية والمفاضلة بين الطرق وإجراء التحليلات الإحصائية.

مجتمع وعينة الدراسة:

عمد الباحث إلى المجتمع الإحصائي، لكونها دراسة محاكاة، حيث تم توليد العينات بحجم (٢٠٠، ٦٠٠، ١٠٠٠) باستخدام البرامج الإحصائية المتخصصة، الحزم الإحصائية (MASS) و(Mvtnorm) التابعة لبرمجية (R)، وبرنامج (SPSS)، و برنامج (WINGEN3) (Excel) ، والتي توفر بيانات وفق ظروف الدراسة الحالية محددة تلائم طبيعة الدراسة، واعتمد الباحث بأن تكون أصغر عينة تمثل خمسة أضعاف المفردة وهي (٢٠٠)، وخمسة عشر أضعاف المفردة وتمثل حجم العينة المتوسط وهي(٦٠٠) كونها الأعلى بناء على ما أوصى به (تيغزة ٢٠١١, ٢٤)، وخمسة وعشرين ضعافا للمفردة تمثل حجم العينة الكبير وهي(١٠٠٠)، علما بأن عدد مفردات الدراسة (٤٠) والتي سبق ذكرها.

عرض وتفسير ومناقشة نتائج الدِّراسَة

١- نتائج الإجابة عن السؤال الأول وتفسيرها:

ISSN: 2537-0464 eISSN: 2537-0472

نص السؤال الاول على" ما تأثير طرق معالجة البيانات المفقودة (تعظيم التوقعات، الانحدار، التعويض المتعدد) على قوة الاختبار الإحصائي والدلالة العملية لمجموعة واحدة؟

للإجابة عن هذا السؤال تم حساب قوة الاختبار الإحصائي بدلالة مربع ايتا ($\eta 2$) لاختبار ت لعينة واحدة ولكل حجم عينة على حِدة في الملفات الثلاثة قبل الفقد ، كما تم حساب قيمة قوة الاختبار الإحصائي بدلالة مربع ايتا للحالات الثلاث من نسب الفقد وفقاً لكل طريقة بالملفات التسعة بعد إجراء التعويض للقيم المفقودة باستخدام البرنامج الإحصائي (SPSS)، حيث تم إعداد ثلاثة ملفات كل ملف يحتوي على احدى هذه النسب من القيم المفقودة، وبعد ذلك تمّ معالجة كل ملف من الملفات الثلاثة بواسطة طرق المعالجة الثلاث (تعظيم التوقعات، الانحدار، التعويض المتعدد) وبناءً على هذه الإجراءات تمّ تكوين تسعة ملفات تبعاً لنسبة فقد البيانات وطريقة معالجتها وهي كالآتي: ثلاثة ملفات تتعلق بطريقة تعظيم التوقعات (EM) (EM10%, EM20%, EM40%) وبالأحجام الثلاث (٢٠٠، ٦٠٠٠، ١٠٠٠) وكذلك ملفات تتعلق بطريقة الانحدار بالنسب الثلاث (Reg10%, Reg20%, Reg40) وملفات تتعلق بطريقة التعويض المتعدد (MI10%, MI20% MI40%,)، ويلخص جدول(٢) النتائج المتعلقة بالمجموعة الواحدة. ثم، تم حساب الدلالة العملية لكل طريقة معالجة ونسبة فقد مع احجام العينات المختلفة.

جدول (٢) ملخص قوة الاختبار الاحصائي لاختبار ت لعينة واحدة وفقاً لحجم العينة ونسبة الفقد وطريقة المعالجة

م	نسبة الفقد	طريقة معالجة الفقد	قوة الاختبار الاحصائي
		حجم العينة ن= ٢٠٠	
١	%٠	قبل الفقد	٠.٩٩٨٠
٢	%١٠	تعظيم التوقعات	٠.٩٤٩٠
٣	%١٠	الانحدار	٠.٨٩٥٠
٤	%١٠	التعويض المتعدد	٠.٩٦٠٠
٥	%٢٠	تعظيم التوقعات	٠.٧٥٢٠
٦	%٢٠	الانحدار	٠.٦١٨٠
٧	%٢٠	التعويض المتعدد	٠.٧٣٥٢
٨	%٤٠	تعظيم التوقعات	٠.٥٣٩٠
٩	%٤٠	الانحدار	٠.٢٤٥٠
١٠	%٤٠	التعويض المتعدد	٠.٥٢٩٠
		حجم العينة ن=٦٠٠	
١	%٠	قبل الفقد	٠.٩٥٩٠
٢	%١٠	تعظيم التوقعات	٠.٩٤٦٠

ISSN: 2537-0464 eISSN: 2537-0472

٣	%١٠	الانحدار	٠.٨٤٢٠
٤	%١٠	التعويض المتعدد	٠.٨٩١٣
٥	%٢٠	تعظيم التوقعات	٠.٩٠١٠
٦	%٢٠	الانحدار	٠.٧٢١٠
٧	%٢٠	التعويض المتعدد	٠.٩٠٦٠
٨	%٤٠	تعظيم التوقعات	٠.٣٣٦٠
٩	%٤٠	الانحدار	٠.٢٢٨٠
١٠	%٤٠	التعويض المتعدد	٠.٢٩٨٠
		حجم العينة ن= ١٠٠٠	
١	%٠	قبل الفقد	٠.٩٨١٠
٢	%١٠	تعظيم التوقعات	٠.٩٤٣٠
٣	%١٠	الانحدار	٠.٨٧١٠
٤	%١٠	التعويض المتعدد	٠.٩٦٤٠
٥	%٢٠	تعظيم التوقعات	٠.٨٥٦٠
٦	%٢٠	الانحدار	٠.٦٢٣٥
٧	%٢٠	التعويض المتعدد	٠.٧١٥٨
٨	%٤٠	تعظيم التوقعات	٠.٥١٢٢
٩	%٤٠	الانحدار	٠.٣٢٩٠
١٠	%٤٠	التعويض المتعدد	٠.٥٧٣٠

تتبين من نتائج جدول (٢) وجود فروق ظاهرية بين قيم قوة الاختبار الاحصائي بين الطرق الثلاث وقيمة قوة الاختبار الإحصائي للبيانات قبل الفقد (نموذج البيانات التامة) حيث أنه من خلال النظر بهذه القيم يتبين ما يلي:

- حجم العينة (ن=٢٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي كانت لطريقة التعويض المتعدد بلغت (٠.٩٦) ثم تعظيم التوقعات (٠.٩٥) بينما بلغت لطريقة الانحدار (٠.٩٠)، وعند نسبة الفقد (٢٠%) كانت لتعظيم التوقعات بلغت (٠.٧٥) ثم لتعويض المتعدد (٠.٧٤) بينما بلغت لطريقة الانحدار (٠.٦٢)، وعند نسبة الفقد (٤٠%) كانت لتعظيم التوقعات بلغت (٠.٥٤) ثم للتعويض المتعدد (٠.٥٣) بينما بلغت لطريقة الانحدار (٠.٢٥)، كما يُلاحظ انخفاض قيمة قوة الاختبار الإحصائي كلما ارتفعت نسبة الفقد.
- حجم العينة (ن=٦٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي كانت لطريقة تعظيم التوقعات بلغت (٠.٩٥) ثم التعويض المتعدد (٠.٨٩) بينما بلغت لطريقة الانحدار (٠.٨٤)، وعند نسبة الفقد (٢٠%) كانت للتعويض المتعدد بلغت (٠.٩١) ثم لتعظيم التوقعات (٠.٩٠) بينما بلغت لطريقة الانحدار (٠.٧٢)، وعند نسبة الفقد (٤٠%) كانت لتعظيم التوقعات بلغت (٠.٣٤) ثم

ISSN: 2537-0464 eISSN: 2537-0472

للتعويض المتعدد (٠.٣٠) بينما بلغت لطريقة الانحدار (٠.٢٣)، كما يُلاحظ انخفاض بعض قيم قوة الاختبار الإحصائي كلما ارتفعت نسبة الفقد وخاصّة لطريقة الانحدار عند نسبة الفقد(٢٠%) ولجميع الطرق عند نسبة الفقد (٤٠%)، كما يُلاحظ انخفاض قيمة قوة الاختبار الإحصائي كلما ارتفعت نسبة الفقد.

- حجم العينة (١٠٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي كانت لطريقة التعويض المتعدد بلغت (٠.٩٦) ثم تعظيم التوقعات (٠.٩٤) بينما بلغت لطريقة الانحدار (٠.٨٧)، وعند نسبة الفقد (٢٠%) كانت لتعظيم التوقعات بلغت (٠.٨٦) ثم للتعويض المتعدد (٠.٧٢) بينما بلغت لطريقة الانحدار (٠.٦٢)، وعند نسبة الفقد (٤٠%)، كانت للتعويض المتعدد (٠.٥٧) ولتعظيم التوقعات (٠.٥١) وللانحدار (٠.٣٣) كما يُلاحظ انخفاض قيم قوة الاختبار الإحصائي كلما ارتفعت نسبة الفقد.

ومن أجل تحديد التأثيرات الرئيسية والتفاعل بينهما في التصميم العاملي، كان من الضروري تحديد نوع الاختبار الاحصائي (F) ومن أجل ذلك تمّ التحقق من كروية البيانات (Sphericity) من خلال فحص الفرضية الصفرية والتي تنص على أنه " لا توجد فروق دالة إحصائياً عند مستوى الدلالة ($\alpha \leq$ ٠,٠٥) بين تباينات الفروق لقيم مستويات المعالجة المختلفة بالتصميم العاملي (3x3). ولاختبار هذه الفرضية تمّ حساب إحصائي حده انحراف الكروية – مواجليز- (Mauchly's) واختبارات (Є)(Epslion) المصححة له (– Greenhouse- Geisser Lower bound) حيث تبين عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (٠.٠٥) بين تباينات الفروق لمستويات المتغيرات المختلفة، حيث بلغت قيمة (Mauchly's) (٠.٣٠٠) وإختبار كاي تربيع (٥.٦٩٢) ودلالتها الإحصائية (٠.٣٤٦).

وبالتالي تم الاعتماد على إحصائي (F) تبعاً لاختبار الكروية (Sphericity Assumed) لفحص التأثيرات بين المعالجات المختلفة، ويوضح جدول (٣) نتائج فحص التأثيرات بين المعالجات المختلفة في التصميم العاملي للمجموعة الأولى.

جدول (٣) تأثير طريقة معالجة البيانات ونسبة الفقد والتفاعل بينهما لقوة الاختبار الاحصائي لاختبار ت لعينة واحدة بالعيّنات الثلاث

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة F	الدلالة الإحصائية	مربع ايتا الجزئي
طريقة المعالجة	٠.٦٨١	٣	٠.٢٢٧	٨٣.٨٣١	٠.٠٠٠	٠.٩٣٣
الخطأ	٠.٠٤٩	١٨	٠.٠٠٣			
نسبة الفقد	٠.٩٥٥	٢	٠.٤٧٧	٣٩.٥٤٣	٠.٠٠٠	٠.٩٢٩
الخطأ	٠.٠٧٢	٦	٠.٠١٢			

ISSN: 2537-0464 eISSN: 2537-0472

تفاعل طريقة المعالجة بنسبة الفقد	٠.٣٣٧	٦	٠.٠٥٦	٢٠.٧٣٦	٠.٠٠٠	٠.٨٧٤
الخطأ	٠.٠١٨	٦	٠.٠٠٣			

يتبين من جدول (٣) وجود فروق دالة إحصائياً بين متوسطات قيم قوة الاختبار الإحصائي تُعزى لطريقة المعالجة حيث بلغت قيمة ف (٨٣.٨٣١) وكانت دلالتها الاحصائية تقل عن مستوى الدلالة (٠.٠٥)

وكانت درجة تأثيرها (٩٣.٣%) كما تبين وجود فروق دالة إحصائياً لقوة الاختبار الإحصائي تُعزى لنسب القيم المفقودة حيث بلغت قيمة ف (٣٩.٥٤٣) ودلالتها الاحصائية (٠.٠٠٠) بدرجة تأثير عالية بلغت (٩٢.٩%) ووجود أثراً للتفاعل حيث تبين أن قيمة (ف) (٢٠.٧٣٦) وكانت دلالتها الاحصائية (٠.٠٠٠) وهي قيمة تقل عن مستوى الدلالة (٠.٠٥) بدرجة تأثير عالية بلغت (٨٧.٤%)، كما تم حساب الأوساط الحسابية الهامشية والخطأ المعياري لقيم قوة الاختبار الاحصائي تبعاً لمتغيري طريقة المعالجة ونسبة القيم المفقودة كما هو موضح بالجدول (٤):

جدول (٤) الأوساط الحسابية الهامشية والأخطاء المعيارية لقيم قوة الاختبار الاحصائي لمجموعة واحدة تبعاً لمتغيري نسبة القيم المفقودة وطريقة معالجتها

المستويات	المتوسط الحسابي	الخطأ المعياري	فترة الثقة عند ٩٥%	
			الحد الأدنى	الحد الأعلى
كامل البيانات	٠.٩٧٩	٠.٠٠٧	٠.٩٦٣	٠.٩٩٥
تعظيم التوقعات	٠.٧٤٨	٠.٠٢٦	٠.٦٨٥	٠.٨١١
الانحدار	٠.٥٩٧	٠.٠١٦	٠.٥٥٨	٠.٦٣٦
التعويض المتعدد	٠.٧٣٠	٠.٠٣٦	٠.٦٤٣	٠.٨١٨
نسبة الفقد ١٠%	٠.٩٣٣	٠.٠٣٢	٠.٨٥٦	١.٠١١
نسبة الفقد ٢٠%	٠.٨١٤	٠.٠٣٢	٠.٧٣٦	٠.٨٩١
نسبة الفقد ٤٠%	٠.٥٤٤	٠.٠٣٢	٠.٤٦٦	٠.٦٢٢

نسبة الفقد	طريقة المعالجة	المتوسط الحسابي	الخطأ المعياري	الحد الأدنى	الحد الأعلى
نسبة الفقد ١٠%	كامل البيانات	٠.٩٧٩	٠.٠١١	٠.٩٥٢	١.٠٠٧
	تعظيم التوقعات	٠.٩٤٦	٠.٠٤٥	٠.٨٣٧	١.٠٥٥
	الانحدار	٠.٨٦٩	٠.٠٢٨	٠.٨٠١	٠.٩٣٨
	التعويض المتعدد	٠.٩٣٨	٠.٠٦٢	٠.٧٨٧	١.٠٩٠
نسبة الفقد ٢٠%	كامل البيانات	٠.٩٧٩	٠.٠١١	٠.٩٥٢	١.٠٠٧
	تعظيم التوقعات	٠.٨٣٦	٠.٠٤٥	٠.٧٢٧	٠.٩٤٦
	الانحدار	٠.٦٥٤	٠.٠٢٨	٠.٥٨٦	٠.٧٢٢
	التعويض المتعدد	٠.٧٨٦	٠.٠٦٢	٠.٦٣٤	٠.٩٣٧
نسبة الفقد ٤٠%	كامل البيانات	٠.٩٧٩	٠.٠١١	٠.٩٥٢	١.٠٠٧
	تعظيم التوقعات	٠.٤٦٢	٠.٠٤٥	٠.٣٥٣	٠.٥٧٢
	الانحدار	٠.٢٦٧	٠.٠٢٨	٠.١٩٩	٠.٣٣٦

ISSN: 2537-0464 eISSN: 2537-0472

التعويض المتعدد	٠.٤٦٧	٠.٠٦٢	٠.٣١٥	٠.٦١٨

وعند قراءة المتوسطات الحسابية الهامشية الناتجة عن تقاطع مستويات متغيرات الدراسة (طريقة المعالجة، نسبة الفقد، التفاعل بينهما) يُلاحظ وجود فروق ظاهرية لصالح تعظيم التوقعات والتعويض المتعدد مقارنة بطريقة التعويض بالانحدار، كما يُلاحظ وجود فروق لنسب الفقد (١٠%) ثم (٢٠%) مقابل (٤٠%) وانخفاض قيمة قوة الاختبار الاحصائي عند نسبة الفقد عبر جميع طرق المعالجة، وكذلك عند كل مستوى من مستويات نسبة الفقد جاءت الفروق بقوة الاختبار الاحصائي لصالح طريقة تعظيم التوقعات وطريقة التعويض مقابل الانحدار، وللكشف عن دلالة الفروق الملاحظة بين قيم المتوسطات الحسابية الهامشية لقيم قوة الاختبار الاحصائي تبعاً لمتغيري نسبة القيم المفقودة وطرق معالجتها، وقد تم استخدام اختبار بينفروني (Bonferroni) للمقارنات البعدية كما تتبين النتائج في جدول (٥).

جدول (٥) نتائج تعديل اختبار بونيفروني (Bonferroni) للمقارنات البعدية لمتوسطات قوة الاختبار الاحصائي تبعاً لمتغيرات الدراسة

	كامل البيانات	تعظيم التوقعات	الانحدار	التعويض المتعدد
كامل البيانات	-	*٠.٢٣١	*٠.٣٨٢	*٠.٢٤٩
تعظيم التوقعات	-	-	*٠.١٥١	-
الانحدار	-	-	-	-
التعويض المتعدد	-	-	*٠.١٣٣	

	نسبة الفقد ١٠%	نسبة الفقد ٢٠%	نسبة الفقد ٤٠%
نسبة الفقد ١٠%	-	*٠.١١٩	*٠.٣٨٩
نسبة الفقد ٢٠%	-	-	*٠.٢٧٠
نسبة الفقد ٤٠%	-	-	-

يلاحظ من الجدول (٥) وجود فرق ذو دلالة احصائية عند مستوى دلالة ($\alpha \leq$ ٠,٠٥) لمتوسطات قوة الاختبار الإحصائي تبعاً لمتغير طريقة المعالجة، إذ بينت نتائج المقارنة الثنائية لمتوسطات قوة الاختبار الإحصائي وجود فروقاً لصالح طريقة التعويض المتعدد وطريقة تعظيم التوقعات مقابل طريقة التعويض بالانحدار ويمكن الاستنتاج بان طريقة التعويض المتعدد وتعظيم التوقع هما الاكثر فاعلية في تقدير القيم المفقودة وذلك بإعطائهما أعلى متوسط هامشي لقيم قوة الاختبار الإحصائي مقارنة بطريقة التعويض بالانحدار ، كما كانت الفروق بحسب نسبة الفقد لصالح (١٠%) مقابل (٢٠%، ٤٠%) وأيضاً فروقاً لصالح (٢٠%) مقابل (٤٠%) ويمكن الاستنتاج بأنه كلما زادت نسبة الفقد بالبيانات تقل قوة الاختبار الإحصائي وتبتعد عن القيمة الحقيقية (قبل الفقد) ويمكن تفسير ذلك بأنّ قيمة قوة الاختبار

ISSN: 2537-0464 eISSN: 2537-0472

الإحصائي تعتمد على قوة التقدير لمعالم الفقرة حيث كلما كانت المعلومات كاملة عن الفقرة ولم يتم فقدها تساهم في قوة الاختبار الإحصائي فزيادة دقة الفقرة تجعل من الفقرة أكثر فاعلية في إتساق استجابات المفحوصين عليها وينعكس ذلك على قيمة قوة الاختبار الإحصائي بالزيادة.

كما تم حساب الدلالة العملية لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (٢٠٠) على الدلالة العملية وفق معيار كوهن، ويلخص الجدول (٦) النتائج على النحو التالي:

جدول (٦) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (٢٠٠) على الدلالة العملية لاختبار ت

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
%٠	قبل الفقد	٠.٠٩٤
%١٠	تعظيم التوقعات	٠.٠٨١
	الانحدار	٠.٠٧٤
	التعويض المتعدد	٠.٠٨٣
%٢٠	تعظيم التوقعات	٠.٠٥
	الانحدار	٠.٠٣٩
	التعويض المتعدد	٠.٠٧١
%٤٠	تعظيم التوقعات	٠.٠٦٤
	الانحدار	٠.٠٥٩
	التعويض المتعدد	٠.٠٥٦

يتبين من نتائج الجدول (٦) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية عند حجم عينة (٢٠٠) لمجموعة واحدة، وكما هو موضح بالملحق رقم (٢)، وكانت النتائج كما يلي:

يمكن مشاهدة أن كل من نسبة الفقد وطريقة المعالجة تؤثر على الدلالة العملية. بشكل عام كلما زادت نسبة الفقد، كلما انخفضت الدلالة العملية نسبيا. وبالمثل، بعض الطرق أكثر فعالية في التعامل مع البيانات المفقودة من غيرها. على سبيل المثال، طريقة التعويض المتعدد تتأثر بشكل أقل بزيادة نسبة الفقد مقارنة بطريقة الانحدار. ويمكن ملاحظة النتائج حيث البيانات الكاملة تحافظ على قوة الاختبار والدلالة العملية، ونستنتج ما يلي:

- **قبل الفقد (٠% فقد)** :في هذه الحالة، الدلالة العملية هي الأعلى وتبلغ (٠.٠٩٤) هذا يشير إلى أن الدلالة العملية متوسطة عندما لا يوجد فقد للبيانات.

ISSN: 2537-0464 eISSN: 2537-0472

- **تعظيم التوقعات (١٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٨١). هذا يشير إلى أن الدلالة العملية متوسطة، ولكنها أقل قوة مقارنةً بالحالة التي لا يوجد فيها فقد.
- **الانحدار (١٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٧٤). هذا يشير إلى أن الدلالة العملية متوسطة، ولكنها أقل قوة مقارنةً بتعظيم التوقعات في حالة الفقد البالغة (١٠%).
- **التعويض المتعدد (١٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٨٣). هذا يشير إلى أن الدلالة العملية متوسطة، وهي الأعلى بين جميع الطرق في حالة الفقد البالغ لنسبة (١٠%) والأقرب لقيمة الدلالة العلمية قبل الفقد.
- **تعظيم التوقعات (٢٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٥). هذا يشير إلى أن الدلالة العملية ضعيفة.
- **الانحدار (٢٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٣٩). هذا يشير إلى أن الدلالة العملية ضعيفة.
- **التعويض المتعدد (٢٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٧١). هذا يشير إلى أن الدلالة العملية متوسطة القوة.
- **تعظيم التوقعات (٤٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٦٤). هذا يشير إلى أن الدلالة العملية متوسطة.
- **الانحدار (٤٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٥٩). هذا يشير إلى أن الدلالة العملية متوسطة.
- **التعويض المتعدد (٤٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٠٥٦). هذا يشير إلى أن الدلالة العملية متوسطة.

وتم حساب الدلالة العملية لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (٦٠٠) على الدلالة العملية وفق معيار كوهن: ويلخص الجدول (٧) النتائج على النحو التالي:

جدول (٧) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (٦٠٠) على الدلالة العملية لاختبار ت

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
%٠	قبل الفقد	٠.١٤١
%١٠	تعظيم التوقعات	٠.١٤١
	الانحدار	٠.١٢٨
	التعويض المتعدد	٠.١٤٢
%٢٠	تعظيم التوقعات	٠.١٣٦
	الانحدار	٠.١٢٨
	التعويض المتعدد	٠.١٣٦

ISSN: 2537-0464 eISSN: 2537-0472

%٤٠	تعظيم التوقعات	٠.١٤٩
	الانحدار	٠.١٤١
	التعويض المتعدد	٠.١٤٦

يتبين من نتائج الجدول (٧) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية عند حجم عينة (٦٠٠) لمجموعة واحدة، وكما هو موضح بالملحق رقم (٣)، وكانت النتائج كما يلي:

يمكن مشاهدة أن كل من نسبة الفقد وطريقة المعالجة وحجم العينة لم تؤثر على الدلالة العملية حيث تراوحت ما بين المتوسط إلى المرتفع. حيث لوحظ مع زيادة حجم العينة أن الدلالة العملية أفضل في الاستقرار ما بين البيانات الكاملة والبيانات التي تم معالجتها بعد الفقد، للتأثير على الدلالة العلمية حيث ظهرت متقاربة وأعلى من نتائج حجم العينة (٢٠٠)، وبعض الطرق أكثر فعالية في التعامل مع البيانات المفقودة من غيرها. على سبيل المثال، طريقة تعظيم التوقعات والتعويض المتعدد تتأثر بشكل أقل بزيادة نسبة الفقد مقارنة بطريقة الانحدار. إلا أن الفارق ليس كبير بين الطرق المختلفة (تعظيم التوقعات، الانحدار، التعويض المتعدد). مما ساعد في الاعتماد عليها في المعالجة من الناحية العملية. ويمكن تفسير ذلك من الجدول كما يلي:

- **قبل الفقد (٠% فقد) :**في هذه الحالة، الدلالة العملية هي الأعلى وتبلغ (٠.١٤١). هذا يشير إلى أن الدلالة العملية مرتفعة نسبيا عندما لا يوجد فقد.
- **تعظيم التوقعات (١٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.١٤١). هذا يشير إلى أن الدلالة العملية مرتفعة نسبيا، ومساوية بالحالة التي لا يوجد فيها فقد.
- **الانحدار (١٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.١٢٨). هذا يشير إلى أن الدلالة العملية متوسطة، ولكنها أقل من الحالة قبل الفقد ومن تعظيم التوقعات.
- **التعويض المتعدد (١٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.١٤٢). هذا يشير إلى أن الدلالة العملية مرتفعة نسبيا وهي قريبة من القيمة قبل الفقد.
- **تعظيم التوقعات (٢٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.١٣٦). هذا يشير إلى أن الدلالة العملية متوسطة وليست بفارق كبير عن قيمتها قبل الفقد.
- **الانحدار (٢٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.١٢٨). هذا يشير إلى أن الدلالة العملية متوسطة.
- **التعويض المتعدد (٢٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.١٣٦). هذا يشير إلى أن الدلالة العملية متوسطة وتتساوى مع تعظيم التوقعات.
- **تعظيم التوقعات (٤٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.١٤٩). هذا يشير إلى أن الدلالة العملية مرتفعة وهي أعلى حتى من قيمة الدلالة العلمية قبل

ISSN: 2537-0464 eISSN: 2537-0472

الفقد والأعلى بين طريقة المعالجة المختلفة ومع نسبة الفقد العالية لـ (٤٠%)، لكنه ليس بفارق كبير.

- **الانحدار (٤٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.١٤١). هذا يشير إلى أن الدلالة العملية متوسطة ومساوية لقيمتها قبل الفقد.
- **التعويض المتعدد (٤٠% فقد) :**في هذه الحالة الدلالة العملية تبلغ (٠.١٤٦). هذا يشير إلى أن الدلالة العملية متوسطة وقريبة لقيمتها قبل الفقد.

وتم حساب الدلالة العملية لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (١٠٠٠) على الدلالة العملية وفق معيار كوهن: ويلخص الجدول (٨) النتائج على النحو التالي:

جدول (٨) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (١٠٠٠) على الدلالة العملية لاختبار ت

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
%٠	قبل الفقد	٠.٥٨٨
%١٠	تعظيم التوقعات	٠.٥٩٥
	الانحدار	٠.٣٧٩
	التعويض المتعدد	٠.٥٦٣
%٢٠	تعظيم التوقعات	٠.٦٤٩
	الانحدار	٠.٤١٨
	التعويض المتعدد	٠.٧٩٨
%٤٠	تعظيم التوقعات	٠.٢٤٥
	الانحدار	٠.٢٢٩
	التعويض المتعدد	٠.٠٧٣

يتبين من نتائج الجدول (٨) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية، كما هو موضح بالملحق رقم (٤)، كما يلي:

يمكن مشاهدة أن كل من نسبة الفقد وطريقة المعالجة وحجم العينة لم تؤثر على الدلالة العملية. كلما زادت نسبة الفقد، كلما انخفضت الدلالة العملية نسبيا بفارق بسيط. مع ملاحظة أن قيمة الدلالة العلمية قبل الفقد وبعد المعالجة لم تختلف كثيرا. مع حجم العينة (١٠٠٠).

ويمكن رؤية أن الدلالة العملية لم تتأثر بشكل كبير بكل من نسبة الفقد وطريقة المعالجة وحجم العينة. وعندما تكون نسبة الفقد أقل، تكون الدلالة العملية أقرب لقيمتها قبل الفقد، وحتى مع نسب الفقد العالية بقيت القيمة مرتفعة مع حجم العينة (١٠٠٠) وفق معيار كوهن.

ISSN: 2537-0464 eISSN: 2537-0472

ومن الجدول السابق (٨) نلاحظ أن جميع طرق معالجة البيانات المفقودة متقاربة في حالة نسبة الفقد (١٠%، و٢٠%) إلا أنه مع زيادة الفقد تأثرت طريقة التعويض المتعدد بشكل نسبي، ويمكن تفسير ذلك من الجدول كما يلي:

- **في حالة (١٠% فقد) :**تعظيم التوقعات يعطي قيمة مرتفعة جدا للدلالة العملية حيث بلغت (٠.٥٩٥).
- **في حالة (١٠% فقد) :**الانحدار يعطي قيمة مرتفعة للدلالة العملية حيث بلغت (٠.٣٧٩).
- **في حالة (١٠% فقد) :**التعويض المتعدد يعطي قيمة مرتفعة للدلالة العلمية حيث بلغت (٠.٥٦٣).
- **تعظيم التوقعات (٢٠% فقد) :**في هذه الحالة الدلالة العملية تبلغ (٠.٦٤٩). هذا يشير إلى أن الدلالة العملية مرتفعة.
- **الانحدار (٢٠% فقد) :**في هذه الحالة الدلالة العملية تبلغ (٠.٤١٨). هذا يشير إلى أن الدلالة العملية مرتفعة.
- **التعويض المتعدد (٢٠% فقد) :**في هذه الحالة الدلالة العملية تبلغ (٠.٧٩٨). هذا يشير إلى أن الدلالة العملية مرتفعة.
- **تعظيم التوقعات (٤٠% فقد) :**في هذه الحالة الدلالة العملية تبلغ (٠.٢٤٥). هذا يشير إلى أن الدلالة العملية مرتفعة.
- **الانحدار (٤٠% فقد) :**في هذه الحالة، الدلالة العملية تبلغ (٠.٢٢٩). هذا يشير إلى أن الدلالة العملية مرتفعة.
- **التعويض المتعدد (٤٠% فقد) :**في هذه الحالة الدلالة العملية تبلغ (٠.٠٧٣). هذا يشير إلى أن الدلالة العملية ضعيفة.

٢- **نتائج الإجابة عن السؤال الثاني وتفسيرها**

نصّ السؤال الثاني على "ما تأثير طرق معالجة البيانات المفقودة (تعظيم التوقعات، الانحدار، التعويض المتعدد) على قوة الاختبار الإحصائي والدلالة العملية لمجموعتين؟ للإجابة عن هذا السؤال تم حساب قوة الاختبار الإحصائي بدلالة مربع ايتا ($\eta 2$) لاختبار ت لعينتين مستقلتين ويلخص جدول (٩) النتائج.

ISSN: 2537-0464 eISSN: 2537-0472

جدول (٩) ملخص قوة الاختبار الاحصائي لاختبار ت لمجموعتين مستقلتين وفقاً لحجم العينة ونسبة الفقد وطريقة المعالجة

م	نسبة الفقد	طريقة معالجة الفقد	قوة الاختبار الاحصائي
	حجم العينة ن= ٢٠٠		
١	٠%	قبل الفقد	٠.٩٩٩
٢	١٠%	تعظيم التوقعات	٠.٨٩٧
٣	١٠%	الانحدار	٠.٧٦٨
٤	١٠%	التعويض المتعدد	٠.٨٧٦
٥	٢٠%	تعظيم التوقعات	٠.٧٩٧
٦	٢٠%	الانحدار	٠.٥٩٧
٧	٢٠%	التعويض المتعدد	٠.٧٥٢
٨	٤٠%	تعظيم التوقعات	٠.٧٠٨
٩	٤٠%	الانحدار	٠.٥٥٨
١٠	٤٠%	التعويض المتعدد	٠.٦٥٨
	حجم العينة ن=٦٠٠		
١	٠%	قبل الفقد	٠.٩٨٠
٢	١٠%	تعظيم التوقعات	٠.٨١٠
٣	١٠%	الانحدار	٠.٤١٠
٤	١٠%	التعويض المتعدد	٠.٧١٠
٥	٢٠%	تعظيم التوقعات	٠.٦٣٠
٦	٢٠%	الانحدار	٠.٤٠٠
٧	٢٠%	التعويض المتعدد	٠.٢٤٠
٨	٤٠%	تعظيم التوقعات	٠.٢٢٠
٩	٤٠%	الانحدار	٠.١٧٠
١٠	٤٠%	التعويض المتعدد	٠.٤٦٩
	حجم العينة ن= ١٠٠٠		
١	٠%	قبل الفقد	٠.٩٩٨
٢	١٠%	تعظيم التوقعات	٠.٧٨٨
٣	١٠%	الانحدار	٠.٦٦٨
٤	١٠%	التعويض المتعدد	٠.٧٥٨
٥	٢٠%	تعظيم التوقعات	٠.٣٢٧
٦	٢٠%	الانحدار	٠.١٧٧
٧	٢٠%	التعويض المتعدد	٠.٣٥٨
٨	٤٠%	تعظيم التوقعات	٠.٢٢٧
٩	٤٠%	الانحدار	٠.١٠٦
١٠	٤٠%	التعويض المتعدد	٠.١٩٨

ISSN: 2537-0464 eISSN: 2537-0472

تتبين من نتائج جدول (٩) وجود فروق ظاهرية بين قيم قوة الاختبار الاحصائي لاختبار ت لمجموعتين مستقلتين بين الطرق الثلاث وقيمة قوة الاختبار الإحصائي للبيانات قبل الفقد (نموذج البيانات التامة) حيث أنه من خلال النظر بهذه القيم يتبين ما يلي:

- حجم العينة (ن=٢٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي بالبيانات الكاملة كانت لطريقة تعظيم التوقعات بلغت (٠.٩٠) ثم التعويض المتعدد (٠.٨٨) بينما بلغت لطريقة الانحدار (٠.٧٨)، وعند نسبة الفقد (٢٠%) كانت لتعظيم التوقعات بلغت (٠.٨٠) ثم لتعويض المتعدد (٠.٧٥) بينما بلغت لطريقة الانحدار (٠.٦٠)، وعند نسبة الفقد (٤٠%) كانت لتعظيم التوقعات بلغت (٠.٧١) ثم للتعويض المتعدد (٠.٦٦) بينما بلغت لطريقة الانحدار (٠.٥٦) ، كما يُلاحظ انخفاض قيمة قوة الاختبار الإحصائي كلما ارتفعت نسبة الفقد.
- حجم العينة (ن=٦٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي كانت لطريقة تعظيم التوقعات بلغت (٠.٨١) ثم التعويض المتعدد (٠.٧١) بينما بلغت لطريقة الانحدار (٠.٤١)، وعند نسبة الفقد (٢٠%) كانت لتعظيم التوقعات بلغت (٠.٦٣) ثم للانحدار (٠.٤٠) بينما بلغت لطريقة التعويض المتعدد (٠.٢٤)، وعند نسبة الفقد (٤٠%) كانت للتعويض المتعدد بلغت (٠.٤٧) ثم لتعظيم التوقعات (٠.٢٢) بينما بلغت لطريقة الانحدار (٠.١٧)، كما يُلاحظ انخفاض قيم قوة الاختبار الإحصائي كلما ارتفعت نسبة الفقد.
- حجم العينة (١٠٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي بالبيانات الكاملة كانت لطريقة تعظيم التوقعات بلغت (٠.٧٩) ثم التعويض المتعدد (٠.٧٦) بينما بلغت لطريقة الانحدار (٠.٦٧)، وعند نسبة الفقد (٢٠%) كانت للتعويض المتعدد بلغت (٠.٣٦) ثم لتعظيم التوقعات (٠.٣٣) بينما بلغت لطريقة الانحدار (٠.١٨)، وعند نسبة (٤٠%) كانت لتعظيم التوقعات (٠.٢٣) ثم للتعويض المتعدد (٠.٢٠) وللانحدار (٠.١١) كما يُلاحظ انخفاض قيم قوة الاختبار الإحصائي كلما ارتفعت نسبة الفقد.

جدول (١٠) تأثير طريقة معالجة البيانات ونسبة الفقد والتفاعل بينهما لقوة الاختبار الاحصائي لاختبار ت لعينتين مستقلتين بالعيّنات الثلاث

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة F	الدلالة الإحصائية	مربع ايتا الجزئي
طريقة المعالجة	١.٥٩٥	٣	٠.٥٣٢	٣٣.٦	٠.٠٠٠	٠.٨٤٩
الخطأ	٠.٢٨٤	١٨	٠.٠١٦			
نسبة الفقد	٠.٥٠٢	٢	٠.٢٥١	٢.٨٢	٠.١٣٧	٠.٤٨٥

ISSN: 2537-0464 eISSN: 2537-0472

الخطأ	٠.٥٣٥	٦	٠.٠٨٩			
تفاعل طريقة المعالجة	٠.٢٠٠	٦	٠.٠٣٣	٢.١١	٠.١٠٣	٠.٤١٣
الخطأ	٠.١٢٢	٦	٠.٠٢٠			

يتبين من جدول (١٠) وجود فروق دالة إحصائياً بين متوسطات قيم قوة الاختبار الإحصائي لاختبار ت لمجموعتين تُعزى لطريقة المعالجة حيث بلغت قيمة ف (٣٣.٦٦٢) وكانت دلالتها الاحصائية تقل عن مستوى الدلالة (٠.٠٥) وكانت درجة تأثيرها (٨٤.٩%) بينما لم يتبين وجود فروق دالة إحصائياً لقوة الاختبار الإحصائي لاختبار ت لمجموعتين مستقلتين تُعزى لنسب القيم المفقودة حيث بلغت قيمة ف (٢.٨٢٠) ودلالتها الاحصائية (٠.١٣٧) بدرجة تأثير متوسطة بلغت (٤٨.٥%) ولم يتبين أثر للتفاعل حيث تبين أن قيمة (ف) (٢.١١٠) وكانت دلالتها الاحصائية (٠.١٠٣) وهي قيمة تزيد عن مستوى الدلالة (٠.٠٥) بدرجة تأثير متوسطة بلغت (٤١.٣%).

ووفقًا لمعيار كوهن، تم تصنيف حجم التأثير على النحو الذي يمكن تفسير حجم التأثير للمتغيرات في الجدول (١٠) على النحو التالي:

طريقة المعالجة: حجم التأثير هو (٠.٨٤٩)، وهذا يعتبر حجم تأثير كبير جدا وفقًا لمعيار كوهن ونسبة الفقد: حجم التأثير هو (٠.٤٨٥)، وهذا يعتبر حجم تأثير متوسط وفقًا لمعيار كوهن. وتفاعل طريقة المعالجة بنسبة الفقد: حجم التأثير كان (٠.٤١٣)، وهذا يعتبر حجم تأثير متوسط وفقًا لمعيار كوهن. كما يمكن ملاحظة التناسب طردي ما بين قوة الاختبار وحجم التأثير.

كما تم حساب الأوساط الحسابية الهامشية والخطأ المعياري لقيم قوة الاختبار الاحصائي تبعاً لمتغيري طريقة المعالجة ونسبة القيم المفقودة كما بالجدول (١١).

جدول (١١) الأوساط الحسابية الهامشية والأخطاء المعيارية لقيم قوة الاختبار الاحصائي لمجموعتين (اختبار ت لعينتين مستقلتين) تبعاً لمتغيري نسبة القيم المفقودة وطريقة معالجتها

المستويات	المتوسط الحسابي	الخطأ المعياري	فترة الثقة عند ٩٥%	
			الحد الأدنى	الحد الأعلى
كامل البيانات	٠.٩٩٢	٠.٠٠٤	٠.٩٨٤	١.٠٠١
تعظيم التوقعات	٠.٦٠٠	٠.٠٧٢	٠.٤٢٥	٠.٧٧٦
الانحدار	٠.٤٢٨	٠.٠٧٢	٠.٢٥٣	٠.٦٠٣
التعويض المتعدد	٠.٥٥٨	٠.٠٧٠	٠.٣٨٦	٠.٧٢٩
نسبة الفقد ١٠%	٠.٨٠٥	٠.٠٨٦	٠.٥٩٤	١.٠١٦
نسبة الفقد ٢٠%	٠.٦٠٥	٠.٠٨٦	٠.٣٩٤	٠.٨١٥
نسبة الفقد ٤٠%	٠.٥٢٤	٠.٠٨٦	٠.٣١٣	٠.٧٣٥
نسبة الفقد	**طريقة المعالجة**	**المتوسط**	**الخطأ**	**الحد الأدنى**	**الحد الأعلى**

ISSN: 2537-0464 eISSN: 2537-0472

المستويات		المتوسط الحسابي	الخطأ المعياري	فترة الثقة عند ٩٥%	
				الحد الأدنى	الحد الأعلى
		الحسابي	المعياري		
نسبة الفقد ١٠%	كامل البيانات	٠.٩٩٢	٠.٠٠٦	٠.٩٧٧	١.٠٠٧
	تعظيم التوقعات	٠.٨٣٢	٠.١٢٤	٠.٥٢٨	١.١٣٥
	الانحدار	٠.٦١٥	٠.١٢٤	٠.٣١٢	٠.٩١٨
	التعويض المتعدد	٠.٧٨١	٠.١٢١	٠.٤٨٤	١.٠٧٨
نسبة الفقد ٢٠%	كامل البيانات	٠.٩٩٢	٠.٠٠٦	٠.٩٧٧	١.٠٠٧
	تعظيم التوقعات	٠.٥٨٥	٠.١٢٤	٠.٢٨١	٠.٨٨٨
	الانحدار	٠.٣٩١	٠.١٢٤	٠.٠٨٨	٠.٦٩٤
	التعويض المتعدد	٠.٤٥٠	٠.١٢١	٠.١٥٣	٠.٧٤٧
نسبة الفقد ٤٠%	كامل البيانات	٠.٩٩٢	٠.٠٠٦	٠.٩٧٧	١.٠٠٧
	تعظيم التوقعات	٠.٣٨٥	٠.١٢٤	٠.٠٨٢	٠.٦٨٨
	الانحدار	٠.٢٧٨	٠.١٢٤	-٠.٠٢٥	٠.٥٨١
	التعويض المتعدد	٠.٤٤٢	٠.١٢١	٠.١٤٥	٠.٧٣٩

وعند قراءة المتوسطات الحسابية الهامشية الناتجة عن تقاطع مستويات متغيرات الدراسة (طريقة المعالجة، نسبة الفقد، التفاعل بينهما) يُلاحظ وجود فروق ظاهرية لصالح تعظيم التوقعات والتعويض المتعدد مقارنة بطريقة التعويض بالانحدار، كما يُلاحظ وجود فروق لنسب الفقد والتفاعل ولكن لا وجود لأثرهما وهي فروق ظاهرية فقط، بينما الفروق كانت لأثر طريقة المعالجة وللكشف عنها تم استخدام اختبار بينفروني (Bonferroni) للمقارنات البعدية كما تبين النتائج ذلك في جدول (١٢).

جدول (١٢) نتائج تعديل اختبار بونيفروني (Bonferroni) للمقارنات البعدية لمتوسطات قوة الاختبار الاحصائي لاختبار ت تبعاً لطريقة المعالجة

	كامل البيانات	تعظيم التوقعات	الانحدار	التعويض المتعدد
كامل البيانات	-	٠.٣٩٢*	٠.١٦٤*	٠.٤٣٥*
تعظيم التوقعات	-	-	٠.١٧٢*	-
الانحدار	-	-	-	-
التعويض المتعدد	-	-	٠.١٣٣*	

يلاحظ من الجدول (١٢) وجود فرق ذو دلالة احصائية عند مستوى دلالة **(٠,٠٥ ≥ α)** لمتوسطات قوة الاختبار الإحصائي تبعاً لمتغير طريقة المعالجة، إذ بينت نتائج

ISSN: 2537-0464 eISSN: 2537-0472

المقارنة الثنائية لمتوسطات قوة الاختبار الإحصائي وجود فروقاً لصالح طريقة التعويض المتعدد وطريقة تعظيم التوقعات مقابل طريقة التعويض بالانحدار ويمكن الاستنتاج بان طريقة التعويض المتعدد وتعظيم التوقع هما الاكثر فاعلية في تقدير القيم المفقودة وذلك بإعطائها أعلى متوسط هامشي لقوة الاختبار الإحصائي لاختبار ت لعينتين مستقلتين مقارنة بطريقة التعويض بالانحدار والشكل (٨) يوضح التفاعل بين المتغيرين (طريقة المعالجة ونسبة الفقد) حيث يلاحظ الاختلافات بسيطة في نسب قيم الفقد مما يُفسر عدم وجود تفاعل بينهما، وربما تفسر النتيجة العامة بأن الفروق للطريقة لا للنسب أو التفاعل لأنه بطريقة حساب قوة الاختبار الاحصائي لمجموعتين مستقلتين يتساوى فيهما التباين وتستطيع الطريقة المستخدمة بمعالجة القيم المفقودة التنبؤ بالقيمة المفقودة بحسب الطريقة والتي لا تتأثر بنسب الفقد لكون التجانس واعتدالية التوزيع الطبيعي للمجموعتين.

وتم حساب الدلالة العملية لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (٢٠٠) على الدلالة العملية وفق معيار كوهن: ويلخص الجدول (١٣) النتائج على النحو التالي:

جدول (١٣) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (٢٠٠) على الدلالة العملية لاختبار ت

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
%٠	قبل الفقد	٠.٠٠١
%١٠	تعظيم التوقعات	٠.٠٠٣
	الانحدار	٠.٠٠٢
	التعويض المتعدد	٠.٠٠٤
%٢٠	تعظيم التوقعات	٠.٠٠٣
	الانحدار	٠.٠٠٣
	التعويض المتعدد	٠.٠٠٨
%٤٠	تعظيم التوقعات	٠.٠٠٢
	الانحدار	٠.٠٠٢
	التعويض المتعدد	٠.٠٠٣

يتبين من نتائج الجدول (١٣) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية عند حجم عينة (٢٠٠) لمجموعتين مستقلتين كما هو موضح بالملحق رقم (٢)، وكانت النتائج كما يلي:

يمكن مشاهدة أن كل من نسبة الفقد وطريقة المعالجة وحجم العينة لم تؤثر على الدلالة العملية، وبشكل عام نلاحظ انخفاض في قيم الدلالة العملية، ويمكن ملاحظة النتائج حيث أن البيانات الكاملة قبل الفقد ضعيفة (٠.٠٠١)، فيما بعد الفقد تتراوح ما بين (٠.٠٠٢ - ٠.٠٠٨) وهي كذلك ضعيفة، مع كافة طرق المعالجة (تعظيم

ISSN: 2537-0464 eISSN: 2537-0472

التوقعات، والانحدار، والتعويض المتعدد)، ومع نسب الفقد المختلفة (١٠%، ٢٠%، ٤٠%) وذلك مع حجم العينة (٢٠٠) وهي العينة التي تمثل خمسة أضاعف المفردة كحد أدنى تم اختيارها من قبل الباحث.كما أشار إلى ذلك تيغزة (٢٠١٢).

وبذلك يرى الباحث أنه يمكن القول بأن عملية التأثير ثابته على الدلالة العملية، قبل المعالجة للبيانات التامة وبعد المعالجة للبيانات المفقودة حيث لم تحدث أي تأثير أو فارق في القيمة وفق معيار كوهن للدلالة العملية. وبذلك يمكن القول بأن عملية المعالجة للبيانات المفقودة تعطي نتائج جيده يمكن الاعتماد عليها.

وتم حساب الدلالة العملية لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (٦٠٠) على الدلالة العملية وفق معيار كوهن: ويلخص الجدول (١٤) النتائج على النحو التالي:

جدول (١٤) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (٦٠٠) على الدلالة العملية لاختبار ت

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
%٠	قبل الفقد	٠.٠٠٠
%١٠	تعظيم التوقعات	٠.٠٠٠
	الانحدار	٠.٠٠٠
	التعويض المتعدد	٠.٠٠٠
%٢٠	تعظيم التوقعات	٠.٠٠٠
	الانحدار	٠.٠٠٠
	التعويض المتعدد	٠.٠٠٠
%٤٠	تعظيم التوقعات	٠.٠٠٠
	الانحدار	٠.٠٠٠
	التعويض المتعدد	٠.٠٠٠

يتبين من نتائج الجدول (١٤) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية عند حجم عينة (٦٠٠) لمجموعتين مستقلتين كما هو موضح بالملحق رقم (٣)، وكانت النتائج كما يلي:

يمكن ملاحظة أن كل من نسبة الفقد وطريقة المعالجة وحجم العينة لم تؤثر على الدلالة العملية. بشكل عام نلاحظ انخفاض في قيم الدلالة العملية. ويمكن ملاحظة النتائج حيث البيانات الكاملة قبل الفقد ضعيفة (٠.٠٠٠)، فيما بعد الفقد أعطت نفس النتيجة (٠.٠٠٠) وهي ضعيفة، مع كافة طرق المعالجة (تعظيم التوقعات، والانحدار، والتعويض المتعدد) ومع نسب الفقد المختلفة (١٠%، ٢٠%، ٤٠%) وذلك مع حجم العينة (٦٠٠) وهي العينة التي تمثل خمسة عشر ضعف للمفردة كحد

ISSN: 2537-0464 eISSN: 2537-0472

متوسط تم اختيارها من قبل الباحث. كما أشار إلى ذلك تيغزة (٢٠١٢) بأنه أفضل حجم مناسب لاختيار العينة.

وبذلك يمكن القول بأن عملية التأثير على الدلالة العملية بقية ثابته قبل المعالجة للبيانات التامة وبعد المعالجة للبيانات المفقودة، حيث لم تحدث أي تأثير أو فارق في القيمة وفق معيار كوهن للدلالة العملية.

وتم حساب الدلالة العملية لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (١٠٠٠) على الدلالة العملية وفق معيار كوهن: ويلخص الجدول (١٥) النتائج على النحو التالي:

جدول (١٥) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (١٠٠٠) على الدلالة العملية لاختبار ت

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
%٠	قبل الفقد	٠.٠٠٢
%١٠	تعظيم التوقعات	٠.٠٠٢
	الانحدار	٠.٠٠٢
	التعويض المتعدد	٠.٠٠٢
%٢٠	تعظيم التوقعات	٠.٠٠٣
	الانحدار	٠.٠٠٣
	التعويض المتعدد	٠.٠٠٢
%٤٠	تعظيم التوقعات	٠.٠٠٣
	الانحدار	٠.٠٠٤
	التعويض المتعدد	٠.٠٠٢

نلاحظ من نتائج الجدول (١٥) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية عند حجم عينة (١٠٠٠) لمجموعتين مستقلتين كما هو موضح بالملحق رقم (٤)، وكانت النتائج كما يلي:

يمكن ملاحظة أن كل من نسبة الفقد وطريقة المعالجة وحجم العينة لم تؤثر على الدلالة العملية. بشكل عام نلاحظ انخفاض في قيم الدلالة العملية. ويمكن ملاحظة النتائج حيث البيانات الكاملة قبل الفقد ضعيفة جدا (٠.٠٠٢)، فيما بعد الفقد تتراوح ما بين (٠.٠٠٢ - ٠.٠٠٤) وهي تكاد أن تكون منعدمة أي ضعيفة جدا، مع كافة طرق المعالجة (تعظيم التوقعات، والانحدار، والتعويض المتعدد) ومع نسب الفقد المختلفة (١٠%، ٢٠%، ٤٠%) وذلك مع حجم العينة (١٠٠٠)، وهي العينة التي تمثل أكبر من خمسة عشرة ضعف للمفردة كحد أعلى تم اختيارها من قبل الباحث.

وبذلك يمكن القول بأن عملية التأثير على الدلالة العملية بقية ثابته قبل المعالجة وبعد المعالجة حيث لم تحدث أي تأثير أو فارق في القيمة وفق معيار كوهن للدلالة العملية، ومن الملاحظ بصورة عامة مع كافة أحجام العينات لاختبار (ت) ظهرت قيم

ISSN: 2537-0464 eISSN: 2537-0472

الدلالة العملية ضعيفة لعينتين مستقلتين، وهذا يعطي مؤشر بأن قبل الفقد وبعد المعالجة كان تأثير المعالجات ثابت فلم تختلف قيم الدلالة العملية، ليعطي دلالة بالاعتماد على طرق المعالجة للبيانات المفقودة باستخدام أي من طرق المعالجة المستخدمة (تعظيم التوقعات، والانحدار، والتعويض المتعدد)، وعند نسب الفقد المختلفة (١٠%، ٢٠%، ٤٠%)، وأحجام العينات (٢٠٠، ٦٠٠، ١٠٠٠)، حيث لم تختلف قيم الدلالة العملية، والتي تساوي تقريبا الصفر، حيث أشار بدوي (٢٠١٨) بأنها دلالة على تساوي المتوسطات بين المجموعات.

٣- نتائج الإجابة عن السؤال الثالث وتفسيرها

نصّ السؤال الثالث على "ما تأثير طرق معالجة البيانات المفقودة (تعظيم التوقعات، والانحدار، والتعويض المتعدد)، على قوة الاختبار الإحصائي لثلاث مجموعات (تحليل التباين الأحادي)؟ للإجابة عن هذا السؤال تم حساب قوة الاختبار الإحصائي بدلالة مربع ايتا ($\eta 2$) لاختبار تحليل التباين لجميع الحالات ويلخص ذلك جدول (١٦) التالي:

جدول (١٦) ملخص قوة الاختبار الاحصائي لتحليل التباين الاحادي وفقاً لحجم العينة ونسبة الفقد وطريقة المعالجة

م	نسبة الفقد	طريقة معالجة الفقد	قوة الاختبار
	حجم العينة ن= ٢٠٠		
١	٠%	قبل الفقد	٠.٩٧٥
٢	١٠%	تعظيم التوقعات	٠.٩٠٦
٣	١٠%	الانحدار	٠.٨٧٩
٤	١٠%	التعويض المتعدد	٠.٨٨٤
٥	٢٠%	تعظيم التوقعات	٠.٨٥١
٦	٢٠%	الانحدار	٠.٧٩٢
٧	٢٠%	التعويض المتعدد	٠.٨٣٦
٨	٤٠%	تعظيم التوقعات	٠.٨٩٣
٩	٤٠%	الانحدار	٠.٨١٦
١٠	٤٠%	التعويض المتعدد	٠.٨٧٥
	حجم العينة ن=٦٠٠		
١	٠%	قبل الفقد	٠.٩٨٧
٢	١٠%	تعظيم التوقعات	٠.٨٨٨
٣	١٠%	الانحدار	٠.٧٨٦
٤	١٠%	التعويض المتعدد	٠.٧١١
٥	٢٠%	تعظيم التوقعات	٠.٧٨٤

ISSN: 2537-0464 eISSN: 2537-0472

٦	%٢٠	الانحدار	٠.٥٥٦
٧	%٢٠	التعويض المتعدد	٠.٧٠١
٨	%٤٠	تعظيم التوقعات	٠.٥٨٧
٩	%٤٠	الانحدار	٠.١٧٦
١٠	%٤٠	التعويض المتعدد	٠.٥٢٥
حجم العينة ن= ١٠٠٠			
١	%٠	قبل الفقد	٠.٩٩١
٢	%١٠	تعظيم التوقعات	٠.٩٠٠
٣	%١٠	الانحدار	٠.٨١٠
٤	%١٠	التعويض المتعدد	٠.٨٦١
٥	%٢٠	تعظيم التوقعات	٠.٨٨١
٦	%٢٠	الانحدار	٠.٧٤٢
٧	%٢٠	التعويض المتعدد	٠.٨٣١
٨	%٤٠	تعظيم التوقعات	٠.٣١٠
٩	%٤٠	الانحدار	٠.١٤٠
١٠	%٤٠	التعويض المتعدد	٠.٢٩٧

يتبين من نتائج جدول (١٦) وجود فروق ظاهرية بين قيم قوة الاختبار الاحصائي بين الطرق الثلاث وقيمة قوة الاختبار الإحصائي للبيانات قبل الفقد (نموذج البيانات التامة) حيث أنه من خلال النظر بهذه القيم يتبين ما يلي:

- حجم العينة (ن=٢٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي لاختبار تحليل التباين كانت لطريقة تعظيم التوقعات بلغت (٠.٩١) ثم التعويض المتعدد (٠.٨٨) بينما بلغت لطريقة الانحدار (٠.٨٧)، وعند نسبة الفقد (٢٠%) كانت لتعظيم التوقعات بلغت (٠.٨٥) ثم لتعويض المتعدد (٠.٨٤) بينما بلغت لطريقة الانحدار (٠.٧٩)، وعند نسبة الفقد (٤٠%) كانت لتعظيم التوقعات بلغت (٠.٨٩) ثم للتعويض المتعدد (٠.٨٧) بينما بلغت لطريقة الانحدار (٠.٨١) ، كما يُلاحظ ثبات قيمة قوة الاختبار الإحصائي باختلاف نسبة الفقد وفروق بسيطة جداً.

- حجم العينة (ن=٦٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي كانت لطريقة تعظيم التوقعات بلغت (٠.٨٩) ثم التعويض المتعدد (٠.٧١) ولطريقة الانحدار (٠.٧٩)، وعند نسبة الفقد (٢٠%) كانت ثم لتعظيم التوقعات بلغت (٠.٧٨) وللمتعدد (٠.٧٠) بينما بلغت لطريقة الانحدار (٠.٥٦)، وعند نسبة الفقد (٤٠%) كانت لتعظيم التوقعات بلغت (٠.٥٩) ثم للتعويض المتعدد (٠.٥٣) بينما بلغت لطريقة الانحدار (٠.١٨) ، كما يُلاحظ انخفاض نسبي لبعض قيم قوة الاختبار الإحصائي كلما ارتفعت نسبة الفقد وخاصّة لطريقة الانحدار .

ISSN: 2537-0464 eISSN: 2537-0472

- حجم العينة (١٠٠٠) يُلاحظ أنه عند نسبة فقد (١٠%) أقرب قيمة لقوة الاختبار الإحصائي كانت لطريقة تعظيم التوقعات بلغت (٠.٩٠) ثم المتعدد (٠.٨٦) بينما بلغت لطريقة الانحدار (٠.٨١)، وعند نسبة الفقد (٢٠%) كانت لتعظيم التوقعات بلغت (٠.٨٨) ثم للتعويض المتعدد (٠.٨٣) بينما بلغت لطريقة الانحدار (٠.٧٤)، وعند نسبة الفقد (٤٠%) ، تعظيم التوقعات(٠.٣١) و للمتعدد (٠.٣٠) وللانحدار(٠.١٤) ويُلاحظ انخفاض قيم قوة الاختبار الإحصائي كلما زاد حجم العينة مع نسبة الفقد وربما لكون أياً من الطرق السابقة عند العينات الكبيرة وزيادة نسبة القيم المفقودة فإن تأثيرها بالتعويض كبير لكونها بيانات لا تمثل القيم الحقيقية ويقل التباين.

ومن أجل اختبار الفرضية الثالثة التي نصّت على " لا توجد فروق دالة إحصائياً عند مستوى دلالة **($\alpha \leq$ ٠,٠٥)** بين متوسطات الحسابية في تقدير قيم قوة الاختبار الإحصائي لثلاث مجموعات تُعزى إلى نسبة القيم المفقودة وطريقة المعالجة والتفاعل بينهما." وذلك للكشف عن دلالة الفروق الإحصائية بين قيم المتوسطات الحسابية لقوة الاختبار الإحصائي وفقاً لأحجامها بالعينات الثلاث التي تُعزى لنسب الفقد وطريقة المعالجة تم استخدام طريقة التحليل العاملي للقياسات المتكررة ثنائي الاتجاه (Two-way repeated measurements method). ومن أجل تحديد التأثيرات الرئيسية والتفاعل بينهما في التصميم العاملي، كان من الضروري تحديد نوع الاختبار الاحصائي (F) ومن أجل ذلك تمّ التحقق من كروية البيانات (Sphericity) من خلال فحص الفرضية الصفرية والتي تنص على أنه " لا توجد فروق دالة إحصائياً عند مستوى الدلالة **($\alpha \leq$ ٠,٠٥)** بين تباينات الفروق لقيم مستويات المعالجة المختلفة بالتصميم العاملي (3x3) ولاختبار هذه الفرضية تمّ حساب إحصائي حده انحراف الكروية – مواجليز- (Mauchly's) واختبارات (Epslion)(Є) المصححة له (Greenhouse- Geisser Lower –bound)حيث تبين عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (٠.٠٥) بين تباينات الفروق لمستويات المتغيرات المختلفة، بلغت قيمة (Mauchly's)(٠.١٠٣) وإختبار كاي تربيع (١٠.٧٣٤) ودلالتها الإحصائية (٠.٠٦٢) وبالتالي تم الاعتماد على إحصائي (F) تبعاً لاختبار الكروية (Sphericity Assumed) لفحص تأثيرات بين المعالجات المختلفة، ويوضح جدول (١٧) نتائج فحص التأثيرات بين المعالجات المختلفة في التصميم العاملي للمجموعة الأولى.

ISSN: 2537-0464 eISSN: 2537-0472

جدول (١٧) تأثير طريقة معالجة البيانات ونسبة الفقد والتفاعل بينهما لقوة الاختبار الاحصائي تحليل التباين الأحادي بالعيّنات الثلاث

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة F	الدلالة الإحصائية	مربع ايتا الجزئي
طريقة المعالجة	٠.٥٩٨	٣	٠.١٩٩	١٦.١٤٢	٠.٠٠٠	٠.٧٢٩
الخطأ	٠.٢٢٢	١٨	٠.٠١٢			
نسبة الفقد	٠.٤١٧	٢	٠.٢٠٨	٢.٦٢١	٠.١٥٢	٠.٤٦٦
الخطأ	٠.٤٧٧	٦	٠.٠٨٠			
تفاعل طريقة المعالجة بنسبة الفقد	٠.١٧٠	٦	٠.٠٢٨	٢.٢٩٢	٠.٠٨١	٠.٤٣٣
الخطأ	٠.١١٩	٦	٠.٠١٩٨			

يتبين من جدول (١٧) وجود فروق دالة إحصائياً بين متوسطات قيم قوة الاختبار الإحصائي لاختبار تحليل التباين تُعزى لطريقة المعالجة حيث بلغت قيمة ف (١٦.١٤٢) وكانت دلالتها الاحصائية تقل عن مستوى الدلالة (٠.٠٥) وكانت درجة تأثيرها (٧٢.٩%) بينما لم يتبين وجود فروق دالة إحصائياً لقوة الاختبار الإحصائي لاختبار تحليل التباين الاحادي تُعزى لنسب القيم المفقودة (٤٦.٦%) ولم يتبين أثر للتفاعل حيث تبين أن قيمة (ف) (٢.٢٩٢) وكانت دلالتها الاحصائية (٠.٠٨١) وهي قيمة تزيد عن مستوى الدلالة (٠.٠٥) بدرجة تأثير متوسطة بلغت (٤٣.٣%) .كما تم حساب الأوساط الحسابية الهامشية والخطأ المعياري لقيم قوة الاختبار الاحصائي تبعاً لمتغيري طريقة المعالجة ونسبة القيم المفقودة كما بالجدول(١٨).

جدول (١٨) الأوساط الحسابية الهامشية والأخطاء المعيارية لقيم قوة الاختبار الاحصائي لتحليل التباين الاحادي تبعاً لمتغيري نسبة القيم المفقودة وطريقة معالجتها

المستويات	المتوسط الحسابي	الخطأ المعياري	فترة الثقة عند ٩٥%	
			الحد الأدنى	الحد الأعلى
كامل البيانات	٠.٩٨٤	٠.٠٠٣	٠.٩٧٨	٠.٩٩١
تعظيم التوقعات	٠.٧٧٨	٠.٠٥٧	٠.٦٣٨	٠.٩١٧
الانحدار	٠.٦٣٣	٠.٠٧٨	٠.٤٤٣	٠.٨٢٣
التعويض المتعدد	٠.٧٢٥	٠.٠٦١	٠.٥٧٦	٠.٨٧٣
نسبة الفقد ١٠%	٠.٨٨٢	٠.٠٨١	٠.٦٨٢	١.٠٨١
نسبة الفقد ٢٠%	٠.٨٢٧	٠.٠٨١	٠.٦٢٨	١.٠٢٦
نسبة الفقد ٤٠%	٠.٦٣١	٠.٠٨١	٠.٤٣٢	٠.٨٣٠

نسبة الفقد	طريقة المعالجة	المتوسط الحسابي	الخطأ المعياري	الحد الأدنى	الحد الأعلى

ISSN: 2537-0464 eISSN: 2537-0472

المستويات		المتوسط الحسابي	الخطأ المعياري	فترة الثقة عند ٩٥%	
				الحد الأدنى	الحد الأعلى
%١٠	كامل البيانات	٠.٩٨٤	٠.٠٠٥	٠.٩٧٣	٠.٩٩٦
	تعظيم التوقعات	٠.٨٩٨	٠.٠٩٩	٠.٦٥٧	١.١٣٩
	الانحدار	٠.٨٢٥	٠.١٣٤	٠.٤٩٦	١.١٥٤
	التعويض المتعدد	٠.٨١٩	٠.١٠٥	٠.٥٦١	١.٠٧٦
%٢٠	كامل البيانات	٠.٩٨٤	٠.٠٠٥	٠.٩٧٣	٠.٩٩٦
	تعظيم التوقعات	٠.٨٣٩	٠.٠٩٩	٠.٥٩٧	١.٠٨٠
	الانحدار	٠.٦٩٧	٠.١٣٤	٠.٣٦٨	١.٠٢٥
	التعويض المتعدد	٠.٧٨٩	٠.١٠٥	٠.٥٣٢	١.٠٤٧
%٤٠	كامل البيانات	٠.٩٨٤	٠.٠٠٥	٠.٩٧٣	٠.٩٩٦
	تعظيم التوقعات	٠.٥٩٧	٠.٠٩٩	٠.٣٥٥	٠.٨٣٨
	الانحدار	٠.٣٧٧	٠.١٣٤	٠.٠٤٩	٠.٧٠٦
	التعويض المتعدد	٠.٥٦٦	٠.١٠٥	٠.٣٠٨	٠.٨٢٣

وعند قراءة المتوسطات الحسابية الهامشية الناتجة عن تقاطع مستويات متغيرات الدراسة (طريقة المعالجة، نسبة الفقد، التفاعل بينهما) يُلاحظ وجود فروق ظاهرية لصالح تعظيم التوقعات والتعويض المتعدد مقارنة بطريقة التعويض بالانحدار، كما يُلاحظ وجود فروق لنسب الفقد والتفاعل ولكن لا وجود لأثرهما وهي فروق ظاهرية فقط، بينما الفروق كانت لأثر طريقة المعالجة وللكشف عنها تم استخدام اختبار بينفروني (Bonferroni) للمقارنات البعدية كما تبين النتائج ذلك في جدول (١٩).

جدول (١٩) نتائج تعديل اختبار بونيفروني (Bonferroni) للمقارنات البعدية لمتوسطات قوة الاختبار الاحصائي لتحليل التباين الاحادي تبعاً لطريقة المعالجة

طريقة المعالجة	كامل البيانات	تعظيم التوقعات	الانحدار	التعويض المتعدد
كامل البيانات	-	-	*٠.٣٥١	*٠.٢٦٠
تعظيم التوقعات	-	-	-	-
الانحدار	-	-	-	-
التعويض المتعدد	-	-	-	

يلاحظ من الجدول (١٩) وجود فرق ذو دلالة احصائية عند مستوى دلالة ($\alpha \leq$ ٠,٠٥) لمتوسطات قوة الاختبار الإحصائي تبعاً لمتغير طريقة المعالجة، لصالح كامل البيانات مقابل التعويض بالانحدار والتعويض المتعدد، ويمكن الاستنتاج بان طريقة الانحدار والتعويض المتعدد هما الأكثر إبتعاداً عن القيمة الحقيقية قبل الفقد وبالتالي تُعد تعظيم التوقعات أفضل منهما لكون الفروق مع كامل البيانات غير

ISSN: 2537-0464 eISSN: 2537-0472

حقيقية والشكل (١٠) يوضح التفاعل بين المتغيرين (طريقة المعالجة ونسبة الفقد) لقوة اختبار تحليل التباين حيث يلاحظ الاختلافات بسيطة في نسب قيم الفقد مما يُفسر عدم وجود تفاعل بينهما، وربما تفسر النتيجة العامة بأن الفروق للطريقة لا للنسب أو التفاعل لأنه بطريقة حساب قوة الاختبار الاحصائي للمجموعات الثلاث يتساوى فيهم تجانس التباين وتستطيع الطريقة المستخدمة بمعالجة القيم المفقودة التنبؤ بالقيمة المفقودة بحسب الطريقة والتي لا تتأثر بنسب الفقد لكونه تحقق التجانس واعتدالية التوزيع الطبيعي للمجموعات الثلاث.

وتم حساب الدلالة العملية، لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (٢٠٠) على الدلالة العملية وفق معيار كوهن: ويلخص الجدول (٢٠) النتائج على النحو التالي:

جدول (٢٠) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (٢٠٠) على الدلالة العملية لاختبار ف

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
٠ %	قبل الفقد	٠.٠٠٥
10%	تعظيم التوقعات	٠.٠١٤
	الانحدار	٠.٠١١
	التعويض المتعدد	٠.٠١٦
20%	تعظيم التوقعات	٠.٠٠٩
	الانحدار	٠.٠٠٨
	التعويض المتعدد	٠.٠١٤
40%	تعظيم التوقعات	٠.٠٠٧
	الانحدار	٠.٠٠٤
	التعويض المتعدد	٠.٠٠٥

يوضح الجدول (٢٠) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية عند حجم عينة (٢٠٠) لاختبار (ف) كما هو موضح بالملحق (٢)، وكانت النتائج كما يلي:

يمكن مشاهدة أن كل من نسبة الفقد وطريقة المعالجة وحجم العينة لم تؤثر على الدلالة العملية. بشكل عام نلاحظ انخفاض في قيم الدلالة العملية. ويمكن ملاحظة النتائج حيث البيانات الكاملة قبل الفقد ضعيفة (٠.٠٠٥)، فيما بعد الفقد تتراوح ما بين (٠.٠٠٤ - ٠.٠١٦) وهي كذلك ضعيفة، مع كافة طرق المعالجة (تعظيم التوقعات، والانحدار، والتعويض المتعدد) ومع نسب الفقد المختلفة (١٠%، ٢٠%، ٤٠%) وذلك مع حجم العينة (٢٠٠) وهي العينة التي تمثل خمسة أضاعف المفردة كحد أدنى تم اختيارها من قبل الباحث. كما أشار إلى ذلك تيغزة (٢٠١٢).

ISSN: 2537-0464 eISSN: 2537-0472

وبذلك يمكن القول بأن عملية التأثير على الدلالة العملية بقية ثابته قبل المعالجة للبيانات التامة وبعد المعالجة للبيانات المفقودة، حيث لم تحدث أي تأثير أو فارق في القيمة وفق معيار كوهن للدلالة العملية.

وتم حساب الدلالة العملية لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (٦٠٠) على الدلالة العملية وفق معيار كوهن: ويلخص الجدول (٢١) النتائج على النحو التالي:

جدول (٢١) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (٦٠٠) على الدلالة العملية لاختبار ف

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
%٠	قبل الفقد	٠.٠١٣
%١٠	تعظيم التوقعات	٠.٠١٢
	الانحدار	٠.٠١٤
	التعويض المتعدد	٠.٠٠٩
%٢٠	تعظيم التوقعات	٠.٠١٦
	الانحدار	٠.٠١٤
	التعويض المتعدد	٠.٠١٩
%٤٠	تعظيم التوقعات	٠.٠١٣
	الانحدار	٠.٠١٤
	التعويض المتعدد	٠.٠١٥

توضح نتائج الجدول (٢١) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية عند حجم عينة (٦٠٠) لاختبار (ف) كما هو موضح بالملحق رقم (٣)، وكانت النتائج كما يلي:

يمكن مشاهدة أن كل من نسبة الفقد وطريقة المعالجة وحجم العينة لم تؤثر على الدلالة العملية. بشكل عام نلاحظ انخفاض في قيم الدلالة العملية. ويمكن ملاحظة النتائج حيث البيانات الكاملة قبل الفقد ضعيفة جدا (٠.٠١٣)، فيما بعد الفقد تتراوح ما بين (٠.٠٠٩ - ٠.٠١٩) وهي تكاد أن تكون منعدمة أي ضعيفة جدا، مع كافة طرق المعالجة (تعظيم التوقعات، الانحدار، والتعويض المتعدد) ومع نسب الفقد المختلفة (١٠%، ٢٠%، ٤٠%) وذلك مع حجم العينة (٦٠٠) وهي العينة التي تمثل خمسة عشر ضعف للمفردة كحد متوسط تم اختيارها من قبل الباحث. كما أشار إلى ذلك تيغزة (٢٠١٢) بأنه أفضل حجم مناسب لاختيار العينة.

وبذلك يمكن القول بأن عملية التأثير على الدلالة العملية بقية ثابته قبل المعالجة للبيانات التامة وبعد المعالجة للبيانات المفقودة، حيث لم تحدث أي تأثير أو فارق في القيمة وفق معيار كوهن للدلالة العملية.

ISSN: 2537-0464 eISSN: 2537-0472

وتم حساب الدلالة العملية لمعرفة تأثير طريقة المعالجة ونسبة الفقد عند حجم العينة (١٠٠٠) على الدلالة العملية وفق معيار كوهن: ويلخص الجدول (٢٢) النتائج على النحو التالي:

جدول (٢٢) نتائج تأثير طريقة المعالجة ونسبة الفقد وحجم العينة (١٠٠٠) على الدلالة العملية لاختبار ف

نسبة الفقد	طريقة معالجة الفقد	الدلالة العملية
%٠	قبل الفقد	٠.٠٠٩
%١٠	تعظيم التوقعات	٠.٠١٠
	الانحدار	٠.٠١٠
	التعويض المتعدد	٠.٠٠٩
%٢٠	تعظيم التوقعات	٠.٠٠٩
	الانحدار	٠.٠٠٨
	التعويض المتعدد	٠.٠٠٩
%٤٠	تعظيم التوقعات	٠.٠١٠
	الانحدار	٠.٠١٠
	التعويض المتعدد	٠.٠١٣

نلاحظ من نتائج الجدول (٢٢) تأثير طريقة المعالجة ونسبة الفقد على الدلالة العملية عند حجم عينة (١٠٠٠) لاختبار (ف) كما هو موضح بالملحق رقم (٤)، وكانت النتائج كما يلي:

يمكن ملاحظة أن كل من نسبة الفقد وطريقة المعالجة وحجم العينة لم تؤثر على الدلالة العملية. بشكل عام نلاحظ انخفاض في قيم الدلالة العملية. ويمكن ملاحظة النتائج حيث البيانات الكاملة قبل الفقد ضعيفة جدا (٠.٠٠٩)، فيما بعد الفقد تراوحت ما بين (٠.٠٠٨ - ٠.٠١٣) وهي ضعيفة جدا، مع كافة طرق المعالجة (تعظيم التوقعات، والانحدار، والتعويض المتعدد) ومع نسب الفقد المختلفة (١٠%، ٢٠%، ٤٠%) وذلك مع حجم العينة (١٠٠٠)، وهي العينة التي تمثل أكبر من خمسة عشرة ضعف للمفردة كحد أعلى تم اختيارها من قبل الباحث.

ومن الملاحظ بصورة عامة مع كافة أحجام العينات لاختبار (ف) ظهرت قيم الدلالة العملية ضعيفة، ولكن ذلك يعطي مؤشر بأن قبل الفقد للبيانات الكاملة وبعد المعالجة لها، كانت قيم التأثير ثابتة فلم تختلف قيمة الدلالة العملية، وهذا يعطي دلالة بالاعتماد على طرق المعالجة للبيانات المفقودة باستخدام أي من طرق المعالجة المستخدمة من الناحية العملية لـ (تعظيم التوقعات، والانحدار، والتعويض المتعدد)، وعند نسب الفقد المختلفة (١٠%، ٢٠%، ٤٠%)، وأحجام العينات (٢٠٠، ٦٠٠، ١٠٠٠)، حيث بقية قيم الدلالة العملية ثابتة تقريبا صفر، حيث أشار بدوي (٢٠١٨) بأنها دلالة على تساوي المتوسطات بين المجموعات.

ISSN: 2537-0464 eISSN: 2537-0472

التعليق على النتائج:

بصورة عامة اتفقت نتائج الدراسة مع نتيجة دراسة (Enders , 2004) التي تبين فيها أثر استراتيجيات التعامل مع البيانات المفقودة (خوارزمية تعظيم التوقعاتEM) أعطت أقل تحيزاً، ودراسة (Hedden, etal, 2008) التي تبين فيها كفاءة (التعويض المتعدد) وفاعليتها على قوة الاختبار الاحصائي باختبار ت وتحليل التباين الأحادي، كما اتفقت مع نتائج دراسة فينج (Finch,2008) التي تبين كفاءة طريقة تعظيم التوقعات، وطريقة التعويض المتعدد، وأنّ الطريقتين أعطت أخطاء معيارية منخفضة وايضاً اتفقت مع دراسة أليسون (Allison, 2009) التي تبين فيها طرق حساب القيم التعويضية في آلية الفقد (MAR) تمتلك أخطاء معيارية أقل، واتفقت أيضاً مع نتيجة دراسة جاريت(Garrett,2009) التي كشفت عن أثر طريقة التعويض المتعدد من طرق التعويض بالبيانات المفقودة على قوة الاختبار الإحصائي وانخفاض نسبي في القوة بزيادة نسبة البيانات المفقودة، وايضاً اتفقت مع نتيجة دراسة حسين (٢٠١٢) التي تبين فيها فاعلية طريقة تعظيم التوقعات وطريقة التعويض بالانحدار، ودراسة (Finch,2011a) التي تبين فيها أثر طريقة التعويض المتعدد مقارنة بطريقة الحذف عبر حجم العينة المختلفة وبنسب فقد للاستجابات.

وأيضاً اتفقت مع نتيجة دراسة (Finch,2011A,b) والتي تبين فيها أثر طريقة التعويض المتعدد وطريقة التعويض باستخدام الانحدار العشوائي على معدلات الخطأ من النوع الاول والقوة الاحصائية ودراسة (Gemici, etal,2012) التي تبين فيها فاعلية طريقة التعويض المتعدد (MI) تُساعد في التقليل من مشكلة التحيز كما تقلل من الخطأ المعياري للتقدير مقارنة بالطرق الأخرى.

كما اتفقت مع نتيجة دراسة الزعبي (٢٠١٣) التي كشفت عن أثر نسبة البيانات المفقودة وطريقة التعويض عنها في دقة تقدير معالم الفقرات والأفراد، بنسبة (30% 20% 15% 5%)، (تعظيم التوقعات، والقيم التعويضية المتعددة) وأظهرت نتائج الدراسة وجود أثر دال إحصائيا يعزى لكل من: متغير بطريقة التعويض لصالح طريقة حساب قيم تعويضية متعددة. بأفضلية نسبة فقد (5%)، وللتفاعل بين متغيري نسبة الفقد وطريقة معالجتها بأفضلية طريقة حساب قيم تعويضية متعددة عندما كانت نسبة الفقد (5%) وايضاً اتفقت مع نتائج دراسة هيبة (٢٠١٣) التي كشفت عن تأثير طرق معالجة البيانات المفقودة (التعويض المتعدد وطريقة تعظيم التوقعات، وطريقة الانحدار) وبتوزيع بيانات مفقودة بنسب (١٠%، ٢٠%، ٤٠%) على جميع الملفات المولدة حيث ظهرت أفضلية طفيفة جدا لطريقة أقصى توقع عند نسبة فقد للبيانات (٤٠%) بينما عدم اختلاف الطرق الثلاثة فى تأثيرها على صدق البنية عند نسبة فقد ١٠% لجميع احجام العينات بينما في حالة نسب الفقد (٢٠%، ٤٠%) أظهرت

ISSN: 2537-0464 eISSN: 2537-0472

أفضلية واضحة لطريقة تحليل الانحدار في عدم تأثيرها على صدق البنية، واتفقت أيضاً مع نتائج دراسة الرحيل والدرابسة (٢٠١٤) التي تبين فيها أثر طريقة معالجة القيم المفقودة، طريقة تعظيم التوقعات، وطريقة التعويضية المتعدد، وتبين وجود اختلاف في دقة تقدير المعلمات ولصالح طريقة التعويض المتعدد (MI). وفروقاً للتفاعل بين متغيري طريقة التقدير وطريقة المعالجة كما اتفقت مع دراسة (Schmitt, etal ,2015) ودراسة (de Gil, 2015) ودراسة (Mamun, etal, 2016). ودراسة (De Silva, etal.2017) ودراسة (Fu, etal, 2021) و (Goretzko,2022) ودراسة (Adjei, Rose, 2022) ودراسة اللصاصمة (٢٠١٦) ودراسة المجالي (٢٠١٧) حيث تبين فيها فاعلية طريقة خوارزمية تعظيم التوقع والتعويض المتعدد على معدلات الخطأ من النوع الاول والقوة الاحصائية عند نسب فقد الاستجابات (١٠%، ٢٠%،٤٠%) وأحجام مختلفة.

بينما اختلفت جزئياً مع نتيجة دراسة العتيبي (٢٠١٨) التي تبين فيها عدم كفاءة نماذج الانحدار المتعدد وفق حالة نسب الفقد المتوسطة والعالية بالنسبة لحجوم العينات الصغيرة، وفي حالة نسب الفقد العالية بالنسبة لحجوم العينات الكبيرة، كذلك طريقة تعظيم التوقع هي الأفضل في تقدير (مستوى الدلالة الإحصائي) في جميع حجوم العينات.

توصيات الدِّراسَة

في ضوء ما أسفرت عنه الدِّراسَة من نتائج فإن الباحث يوصي بما يلي:

١- استخدام طريقة تعظيم التوقعات والتعويض المتعدد بالدراسات التربوية والنفسية.

٢- عدم الاعتماد على طريقة التعويض بالانحدار، عندما تكون نسبة الفقد ٢٠% فأكثر بالدراسات التربوية والنفسية.

٣- التأكد من استيفاء افتراضات استخدام طرق التعويض قبل الاعتماد عليها.

٤- عدم الاعتماد على طريقة واحدة بعينها للتعامل مع القيم المفقودة وإنما بحسب نوع الفقد والظروف التي أدت الى فقد البيانات حيث يسهم ذلك في تحديد نمط ملأ البيانات والتعامل معها وفق طريقة معالجة البيانات.

٥- لا توجد طريقة مثالية لمعالجة البيانات المفقودة تؤدي إلى الحصول على نتائج دقيقة مماثلة لتلك النتائج فيما لو كانت البيانات تامة؛ لذا على الباحث أنْ يحرص على جمع بيانات تامة قدر الإمكان وعلى الباحث فحص نموذج البيانات المفقودة ومعرفة نمط فقد البيانات وآليتها حتى يتم اختيار أسلوب التعويض المناسب.

ISSN: 2537-0464 eISSN: 2537-0472

قائمة المراجع

أبو بدر سليمان حسن. (٢٠٢٠). *استخدام الأساليب الإحصائية في بحوث العلوم الاجتماعية* دمشق، المركز العربي للأبحاث ودراسة السياسات.

باهي مصطفى و عنان عبدالفتاح. (٢٠٠٦). *الاحصاء التطبيقي باستخدام الحزم الاحصائية*، القاهرة: مكتبة الأنجلو المصرية.

بدوي، عبير علي. (٢٠١٨). مقارنة طرق قياس حجم الأثر لبعض الأساليب الإحصائية مع أحجام عينات مختلفة. *مجلة البحث العلمي في التربية*، ١٩، ٤٣٣-٤٩١.

بني عواد، علي، والمومني، خالد. (٢٠١٨). إدراك أعضاء هيئة التدريس في جامعة الملك فيصل لأهمية طرق التعامل مع البيانات المفقودة في الاستبانات البحثية واختبارات الطلبة .*مجلة جامعة النجاح للأبحاث - العلوم الإنسانية*، ٣٢(٦)، ١٠٣٣ - ١٠٦٥.

تيغزة، محمد بوزيان. (٢٠١١). *التحليل العاملي الاستكشافي والتوكيدي مفاهيمهما ومنهجيتهما بتوظيف حزمة SPSS وليزرل LISREL*، عمان: دار المسيرة للنشر والتوزيع والطباعة.

جودة، محفوظ. (٢٠٠٩). *التحليل الإحصائي المتقدم باستخدامSPSS*، عمّان: دار وائل للنشر.

حسين، علي ناصر. (٢٠١٢). تقدير القيم المفقودة لمتغير الاستجابة في نموذج الانحدار الخطي المتعدد. *العلوم الاقتصادية*، ٨(٣٠)، ٢٣١ - ٢٤٦.

الحيالي، علي درب كسار. (٢٠١٣). استخدام معايير الدقة التنبؤية في تحديد الطريقة المثلى في تقدير القيمة المفقودة: بيانات البحوث الزراعية أنموذجاً .*مجلة العلوم الزراعية العراقية*، ٤٤(٤)، ٥٠٩-٥١٧.

الرحيل، راتب صايل والدرابسة، رياض أحمد. (٢٠١٤). أثر طريقتي التعامل مع القيم المفقودة، وطريقة تقدير القدرة على دقة تقدير معالم الفقرات والأفراد .*المجلة الدولية التربوية المتخصصة*، ٣(٦)، ٢٣-٤٧.

الزعبي، عمر صالح. (٢٠١٣). *أثر نسبة البيانات المفقودة وطريقة التعويض عنها في دقة تقدير معالم الفقرات والأفراد*[رسالة دكتوراه غير منشورة] . جامعة اليرموك، إربد.

الصرايرة، راجي.(٢٠١٨). أثر نسبة القيم المفقودة وطريقة معالجتها في دقة تقدير القيمة القصوى لدالة معلوماتية الفقرة وقيمة قوة الاختبار الإحصائي، *مجلة دراسات، العلوم التربوية*، الجامعة الأردنية، ٤٥(٤) ٤٨٥- ٥٠١.

عبيدات، ذوقان و عبدالحق، كايد وعدس، عبدالرحمن. (٢٠١٠). *البحث العلمي*

ISSN: 2537-0464 eISSN: 2537-0472

مفهومه وأدواته وأساليبه. ط٢ عمان: دار الفكر.
العتيبي، أشرف. (٢٠١٨). تأثير البيانات المفقودة على كفاءة نموذج الانحدار المتعدد وطرق معالجتها] رسالة دكتوراه غير منشورة] . جامعة أم القرى.
العتيبي، أشرف. (٢٠٢٢). المقارنة بين طريقتي التعويض بآخر قيمة ملاحظة وهوت دك لمعالجة البيانات المفقودة في ضوء معيار قوة الاختبار الإحصائي: دراسة محاكاة. *مجلة جامعة الملك خالد للعلوم التربوية*، ٩(٣)، ١١٣ - ١٣٧.
العراقي، عمرو عبدالكريم. (٢٠١٦). *صحافة البيانات خطوات جمع وتحليل البيانات وتصميم الإنفوجرافيك*. القاهرة. العربية للنشر والتوزيع.
علام، صلاح الدين محمود. (٢٠٠٣). *تحليل بيانات البحوث النفسية والتربوية والاجتماعية*، القاهرة: دار الفكر العربي
علام، صلاح الدين محمود. (٢٠١٠). *الأساليب الإحصائية الاستدلالية في تحليل بيانات البحوث النفسية والتربوية والاجتماعية البارمترية واللابارمترية*، القاهرة: دار الفكر العربي.
فهمي، محمد شامل بهاء الدين. (٢٠٠٩). *الاحصاء بلا معاناة: المفاهيم مع التطبيقات باستخدام برنامج SPSS (الجزء الثاني)*. معهد الادارة العامة: الرياض.
كاريان، زافين و دوديويز، إدوارد. (١٩٩٩). *المحاكاة الإحصائية ومحاكاة النظم ونظام المحاكاة متعددة الأغراض* GPSS (ترجمة: سرور علي إبراهيم سرور). الرياض: جامعة الملك سعود.
اللصاصمة، شريفة سليمان. (٢٠١٨). *أثر حجم العينة وطريقة التعامل مع القيم المفقودة على ثبات الاختبار ومعاملات تمييز وصعوبة الفقرات* [رسالة ماجستير غير منشورة]. جامعة مؤتة، مؤتة.
اللصاصمة، عمران إسماعيل. (٢٠١٦). *أثر نسبية القيم المفقودة وطريقة معالجتها في دقة تقدير معالم معادلة الانحداد البسيط] رسالة* ماجستير غير منشورة]. جامعة مؤتة، الكرك.
المجالي، فاطمة جمال. (٢٠١٧). *المقارنة بين فاعلية ثلاث طرق لمعالجة القيم المفقودة في ضوء جودة مطابقة الفقرة للنموذج اللوجستي ثلاثي المعلمة* [رسالة ماجستير غير منشورة]. جامعة مؤتة، مؤتة.
المعهد العربي للتدريب والبحوث الإحصائية. (٢٠٢٣) ورشة عمل بعنوان " معالجة القيم المفقودة في المسوح والتعدادات
https://www.aitrs.org/ar/node/808.
منصور، رشدي فام. (1997). *حجم التأثير- الوجه المكمل للدلالة الإحصائية. المجلة المصرية للدراسات النفسية*. ١٦(٧)، ٥٦-٨٥.
النجار، نبيل. (٢٠١٥). الإحصاء التحليلي مع تطبيقات برمجية SPSS. عمان: دار

ISSN: 2537-0464 eISSN: 2537-0472

ومكتبة الحامد للنشر والتوزيع.

هيبة، محمد أحمد. (٢٠١٣). تأثير طرق معالجة البيانات المفقودة على الخصائص السيكو مترية للمقاييس ذات الاستجابات المتعددة (دراسة إمبيريقية ومحاكاة)، *مجلة جامعة عين شمس للقياس والتقويم*، ٣(٥)، ١-٥٨.

Acock, A. C. (2005). Working with Missing Values. *Journal of Marriage and Family*, 67,1012-1028.

Adjei, R. (2022), *Statistical Challenges and Methods for Missing and Imbalanced Data*" [All Graduate Theses and Dissertations]. 8621. https://digitalcommons.usu.edu/etd/8621

Allison, P. P. (2009). Imputation of Categorical Variables With PROC MI. *Paper presented at the annual meeting of the SAS Users Group International*, San Francisco, CA. 71(3), 933–946. http://www.jstor.org/stable/1555527

Baraldi, A. N., & Enders, C. K. (2010). *An introduction to modern missing data analyses*. Journal of School Psychology, 48, 5-37.

Bori, M. S. (2013). *Dealing with missing data: Key assumptions and methods for applied analysis.* This paper was published in fulfillment of the requirements for PM931 Directed Study in Health Policy and Management under Professor Cindy Christiansen.

Carpenter, J., & Kenward, M. (2012). *Multiple imputation and its application*. John Wiley & Sons https://doi-org.sdl.idm.oclc.org/10.14569/IJACSA.2023.0141049

Data C. C.(2021). *Handling missing values*", Kaggle.com, [Online]. Available: https://www.kaggle.com/rtatman/data-cleaning-challenge-handling-missing-values. [Accessed: 20- 12- 2023].

De Gil, P. R.(2015). *An Empirical Comparison of the Effect of Missing Data on Type I Error and Statistical Power of the Likelihood Ratio Test for Differential Item Functioning:An Item Response Theory Approach using the Graded Response Model*, [Adissertation of Doctor of Philosophy], College of Education, University of South Florida.

De Silva, A. P., Moreno-Betancur, M., De Livera, A. M., Lee, K. J., & Simpson, J. A.(2017). A comparison of multiple imputation methods for handling missing values in longitudinal data in the

ISSN: 2537-0464 eISSN: 2537-0472

presence of a time-varying covariate with a non-linear association with time: a simulation study. *BMC medical research methodology*,17(1),114-218

Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis and the Interpretation of Research Results. Cambridge*: Cambridge.

Enders, C. K. (2004). The impact of missing data on sample reliability estimates: Implications for reliability reporting practices. *Educational and Psychological Measurement,* 64, 419 – 436.

Enders, C. K. (2010*). Applied missing data analysis*. New York: The Guilford Press.

Fan, X.(1999). Statistical Significance and Effect Size: Tow Sides of a Coin. *Paper Presented at the Annual Meeting of the American Evaluation Associatio*n (Orlando, FL, November 3-6, 1999). (ERIC Reproduction Service No.ED435702).

Figueredo, A., Mcknight, P., Mcknight, K., &Sidani, S. (2000). *Multivariate modeling of missing data within and across assessment waves Addiction*, 95, 361-380.

Finch, H. (2008).Estimation of item response theory parameters in the presence of missing data. *Journal of Educational*. 45, 225 –245.

Finch, H. (2011a). The use of multiple imputation for missing data in uniform DIF analysis: Power and type I error rates. *Applied Measurement in Education*, 24, 281-301.

Finch, W. H. (2011b). The impact of missing data on the detection of nonuniform differential item functioning. *Educational and Psychological Measurement,* 71(4), 663-683.

Fu, Y. Hongjian. L. Longlong Lv. (2021). A Comparative Study of Various Methods for Handling Missing Data in UNSODA *Agriculture* 11, no. 8: 727. https://doi.org/10.3390/agriculture11080727

Garrett, P. (2009). *A Monte Carlo study investigating missing data, differential item functioning, and effect size* [Unpublished doctoral dissertation]. Georgia State University.

Gemici, S., Bednarz, A., & Lim, P. (2012). A primer for handling missing values in the analysis of education and training data.

ISSN: 2537-0464 **eISSN: 2537-0472**

International Journal of Training Research, 10(3), 233-250. https://doi.org/10.5172/ijtr.2012.10.3.233

Goretzko D. (2022). Factor Retention in Exploratory Factor Analysis With Missing Data. *Educational and psychological measurement*, 82(3), 444–464. https://doi.org/10.1177/00131644211022031

Gordon Howard R. D. *American Vocational Education Research Association Members*, Perceptions of Statistical significance Tests and other Statistical controversies, 2001.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549 –576.

Graham, J. W. (2012). Missing data analysis and design. New York: springer.

Hedden, S. L., Woolson, R. F., & Malcolm, R. J. (2008). A comparison of missing data methods for hypothesis tests of the treatment effect in substance abuse clinical trials: a Monte-Carlo simulation study. *Substance abuse treatment, prevention, and policy*, 3, 13-14. https://doi.org/10.1186/1747-597X-3-13

Huston, Holly L. (1993). ingfulness, Statistical Significance, Effect Size, and Power Analysis: A General Discussion With Implications for MANOVA. *Paper Presented at the Annual Meeting of the Mid-South Educational Research Association* (New Orleans, LA, November 9-12). (ERIC Reproduction Service NO.ED364608).

Kaiser, J. (2014). Dealing with Missing Values in Data. Journal of Systems Integration, 42-51. https://doi.org/10.20470/jsi.v5i1.178

Kalaycioglu, O., Copas, A., King, M., & Omar, R. Z. (2016). A comparison of multipleimputation methods for handling missing data in repeated measurements observational studies. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(3), 683–706

Karahalios, A., Baglietto, L., Carlin, J. B., English, D. R., & Simpson, J. A. (2012).A review of the reporting and handling of missing

ISSN: 2537-0464 **eISSN: 2537-0472**

data in cohort studies with repeated assessment of exposure measures. BMC medical research methodology,12(1), 96-119

Lee, K. J., & Carlin, J. B. (2010). Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *American journal of epidemiology*, 171(5), 624–632.

Little, R. J., & Rubin, D. B. (2014). *Statistical analysis with missing data* [15] nd edition (Vol. 793). John Wiley & Sons.

Ludlow, L. H., & O'Leary, M. (1999). *Scoring omitted and not-reached items: practical data analysis implications*. Educational and Psychological Measurement, 59(4), 615 −630.

Mamun, A., Zubairi, Y., Hussin, A., & Rana, S . (2016). A comparison of missing data handling methods in linear structural relationship model: evidence from BDHS2007 data. *Electronic Journal of Applied Statistical Analysis.* 122 -133.

Mcknight, P. E., Mcknight, K. M., Sidani, S. & Figuredo, A . J. (2007). *Missingdata a gentle introduction*. New York : The Guilford Press.

Mirzaei, A., Carter, S., A. Patanwala., & Schneider,C. (2021) Missing data in surveys: Key concepts, approaches, and applications", *Research in Social and Administrative Pharmacy,* 2021. Available: 10.1016/j.sapharm.2021.03.009.

Mislevy, R. J., & Wu, P. K. (1988). *Inferring examinee ability when some item responses are missing* (ERIC Document Reproduction Service No. ED 395 017). Princeton, NJ: Educational Testing Service.

Musil, C. M., Warner C. B., Yobas, P. K., and Jones S. L. (2002). A Comparison of Imputation Techniques for Handling Missing Data, *Western Journal of Nursing* 12(2) 9-33.

Onwuegbuzie, A. J.; Levin, J. R.; Leech, N. L.(2003). Do Effect-Size Measures Measure up?: *A Brief Assessment*. A Contemporary Journal1(1) 37-40.

Peugh, J. & Enders, C. (2004). Missing Data in Educational Research: A Review of Reporting Practices and Suggestions for Improvement. *Review of Educational Research - REV EDUC RES*. 74. 525-556. 10.3102/00346543074004525.

ISSN: 2537-0464 **eISSN: 2537-0472**

Pigott, T. D. (2001). A *review of methods for missing data*. Educational Research and Evaluation, 7, 353 –383.

Schafer, J. L., & Graham, J. W. (2002). *Missing data: Ourview of thestate of the art*. Psychological Methods, 7,147–177.

Schmitt P, Mandel J, Guedj M (2015) A Comparison of Six Methods for Missing Data Imputation. *J Biomet Biostat 6*: 224. doi:10.4172/21556180.1000224

Steyn, H. S, Ellis, S. M. (2009). estimating an effect size in one-way multivariate analysis of variance (MANOVA).*multivariate behavioral*

Witta, E.L. (2000). *Effectiveness of four methods of handling missing data using samples from a national database*. Teacher Education Yearbook, 28.

Zhou, Q. (2001). *Missing value imputation methods for parameter estimates and psychometric properties of likert measures.* Thesis submitted to the Faculty of the Graduate School of the University of Maryland in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ISSN: 2537-0464 **eISSN: 2537-0472**