Exploring the Relationship Between Legend Categories and Player Success in Apex Legends

Altaaf Ally (2424551)
School of Computer Science & Applied Mathematics
University of the Witwatersrand
Johannesburg, South Africa

Dr. Branden Ingram (Supervisor)
School of Computer Science & Applied Mathematics
University of the Witwatersrand
Johannesburg, South Africa

Dr. Pravesh Ranchod (Supervisor)
School of Computer Science & Applied Mathematics
University of the Witwatersrand
Johannesburg, South Africa

***Abstract*—*Apex Legends*, a fast-paced battle royale game, offers unique challenges in assessing player performance due to its evolving gameplay and character-specific abilities. This study explores two hypotheses: first, that different Legend categories exhibit distinct performance profiles aligned with their predefined roles, with Assault Legends achieving higher kills per win ratios and kills per match than more defensive or support-oriented Legends; and second, that a strong positive correlation exists between a player's kill count and their overall wins. Using clustering and statistical analysis, we identified distinct performance patterns across Legend categories, validating that offensive Legends excel in kill metrics while tactical roles prioritise efficiency. The positive correlation between Career Kills and Career Wins further underscores the significance of offensive actions for player success. These findings offer valuable insights into player strategies and gameplay balance, providing a foundation for future research as *Apex Legends* continues to evolve.**

***Index Terms*—Apex Legends, Player Performance, Legend Categories, Battle Royale, Data Analytics, Clustering, Correlation Analysis, Regression Analysis, Game Strategy**

I. Introduction

Apex Legends is a team-based battle royale game distinguished by its fast-paced gameplay and a diverse roster of characters, known as Legends, each with unique abilities that emphasise offensive, defensive, or support roles. These character-specific abilities enable players to adopt varied strategies, making performance evaluation more complex than in traditional multiplayer games, where metrics such as kill-death ratio (K/D), win rate, or match ranking are typically sufficient [1]. In *Apex Legends*, a player's effectiveness is shaped by their choice of Legend, adaptability to dynamic in-game situations [2], and role-based contributions [3], presenting unique challenges for assessing individual and team success.

In *Apex Legends*, factors such as character selection, adaptability, and aggression levels strongly impact performance [4], [5], but traditional skill ratings and win/loss metrics often fail to capture these dynamics. Existing systems, including Elo and TrueSkill , primarily focus on overall match outcomes. The Elo system assigns ratings based on match results, while TrueSkill extends this by factoring in both player skill and rating uncertainty, which is useful for team-based games [6]. However, both systems often miss individual contributions and character-specific effects that are essential in team-oriented games like Apex Legends [3], [7]. Therefore, more refined, game-specific metrics are needed. This study addresses this gap by examining two hypotheses related to player performance in *Apex Legends*:

- **Hypothesis 1**: Offensive Legends will achieve higher kills per win and kills per match compared to Legends with defensive or support roles.
- **Hypothesis 2**: There will be a strong positive correlation between a player's kills and their wins, suggesting that kill-based performance is crucial for success.

By analysing datasets of player statistics, this research aims to uncover key performance drivers in *Apex Legends*, focusing on how Legend selection, playstyle aggressiveness, and in-game decision-making impact outcomes [4], [8]. The findings will provide insights useful for players, coaches, and game developers and contribute to the body of knowledge on evaluating player performance in team-based competitive games.

II. Overview of Apex Legends Gameplay

Apex Legends is a popular battle royale game developed by Respawn Entertainment, known for its team-based gameplay and diverse character selection. Players form squads of three and compete against other teams to be the last squad standing on a large, dynamic map. Each player selects a unique ”Legend,” a character with distinct abilities categorised into roles such as Assault, Skirmisher, Recon, Support, and Controller [9].

Gameplay emphasises both survival and combat, with players needing to strategically balance offensive actions (such as engaging opponents) and defensive maneuvers (like positioning and resource management). The game's emphasis on team synergy and character-specific abilities introduces complex dynamics that influence individual and team performance, making it an ideal subject for examining how specific roles and actions contribute to success in competitive settings.

III. Background

A. Multiplayer Online Games and Player Performance

Multiplayer online games, including first-person shooters (FPS), massively multiplayer online role-playing games (MMORPGs), and battle royale games, have gained immense popularity, attracting millions of players worldwide. Understanding player performance in these games is essential for enhancing matchmaking systems, game balancing, and overall player experience [10].

Research on player performance covers a broad range of factors, from team dynamics to technical aspects influencing in-game behavior. For example, Guo et al. [4] examined player workload, win rates, and behavioral patterns across various online games, laying a foundation for analysing performance in games with complex individual and team dynamics, such as battle royale games.

Additionally, team synergy is a critical element. Carrilho et al. [2] used optical tracking data to assess synchronisation in player movements, a concept adaptable to multiplayer games like Apex Legends, where team coordination is key.

B. Player Performance Evaluation Techniques

Traditional evaluation techniques in multiplayer games often rely on skill-based rating systems like Elo or TrueSkill, which mainly consider win/loss outcomes. While widely adopted, these metrics have limitations, particularly in games requiring multifaceted performance, such as teamwork, adaptability, and behavior under dynamic conditions [1].

To address these gaps, researchers have adopted more detailed approaches that include in-game statistics. Fernandez-Navarro et al. [5] explored playstyle effectiveness in elite soccer, highlighting playstyle as a significant factor in team performance. In competitive shooters, Dehpanah et al. [11] developed a behavioral rating system that integrates metrics such as aggressiveness and strategic approach, providing a more comprehensive view of player performance.

C. Battle Royale Games and Player Performance

Battle royale games, such as PlayerUnknown's Battlegrounds (PUBG) and Fortnite, introduced large-scale combat with survival elements, creating unique challenges for performance assessment [12]. In these unpredictable environments, factors like positioning, resource management, and adaptability are crucial for success.

Dehpanah et al. [7] compared traditional rating systems like Elo and TrueSkill within PUBG, finding that these systems struggled to capture the complexity of battle royale gameplay. This highlighted a need for systems that consider the interdependencies between teammates and the genre's fast-paced nature. These findings are particularly relevant for Apex Legends, which incorporates character-specific abilities and emphasises teamwork, adding further layers to individual performance evaluation.

D. Clustering and Statistical Analysis in Player Performance

Clustering and statistical analysis have become essential tools for understanding patterns in player behavior and performance across complex gaming datasets. Clustering allows researchers to group players or characters based on similar characteristics, enabling insights into playstyle trends and role-based performance distinctions [13], [14]. This technique has been successfully applied in gaming research to identify behavioral archetypes, optimise strategies, and enhance game balance [15], [16].

Statistical analysis complements clustering by quantifying relationships between metrics, such as kills and wins, and testing hypotheses about player success. For instance, statistical methods have been used to predict player outcomes and understand the factors contributing to player engagement [17]. This combination of clustering and statistical methods supports a data-driven approach to performance evaluation, allowing for objective insights into how specific Legend roles contribute to team success in Apex Legends.

E. Legend Categories and Abilities

In *Apex Legends*, each Legend possesses three unique abilities—passive, tactical, and ultimate—which align them into five specific categories: Assault, Skirmisher, Recon, Support, and Controller [9]. These roles influence both individual playstyles and team strategies, providing a framework for analysing performance metrics across categories.

1) Assault Legends: Designed for aggressive, high-damage gameplay, Assault Legends enable players to initiate combat effectively. This category includes:

- **Bangalore**, **Fuse**, **Ash**, **Mad Maggie**, and **Ballistic**.

2) Skirmisher Legends: Skirmisher Legends excel in quick engagements and mobility, offering versatility in combat. This category includes:

- **Pathfinder**, **Wraith**, **Octane**, **Revenant**, **Horizon**, **Valkyrie**, and **Alter**.

3) Recon Legends: Recon Legends focus on intelligence gathering and tracking enemy positions, aiding in team situational awareness. This category includes:

- **Bloodhound**, **Crypto**, **Seer**, and **Vantage**.

4) Support Legends: Support Legends provide healing, resources, and revives, sustaining team health during prolonged engagements. This category includes:

- **Gibraltar**, **Lifeline**, **Mirage**, **Loba**, **Newcastle**, and **Conduit**.

5) Controller Legends: Controller Legends specialise in area control, using abilities to lock down locations and limit enemy movement. This category includes:

- **Caustic**, **Wattson**, **Rampart**, and **Catalyst**.

This categorisation of Legends will serve as a foundation for evaluating how different roles contribute to individual performance metrics and overall team success. By analysing performance metrics like Kills per Match and Kills per Win across these roles, we aim to identify trends that differentiate each category's impact on gameplay.

IV. Related Work

A. Evaluating Player Performance in Battle Royale Games

The evaluation of player performance in battle royale games poses unique challenges due to their dynamic, team-based nature. Dehpanah et al. [7] investigated methods for assessing rating systems within these games, highlighting the limitations of traditional models like Elo and TrueSkill in capturing team-oriented and player-specific contributions. Building upon Guo et al. [4]'s foundational research on player workload and behavior in online games, Dehpanah et al. underscored the need for nuanced performance metrics that can account for the interdependencies among players in team-based battle royale games, such as Apex Legends.

Additionally, adaptability to in-game conditions is crucial in dynamic gaming environments. He et al. [18] explored style adaptability in football, a concept potentially applicable to games like Apex Legends, where players continuously adjust to evolving battlefield conditions and shifting team compositions. The ability to modify playstyle in response to game circumstances likely impacts performance, further reinforcing the need for adaptive metrics in evaluating players.

B. Character-Specific Performance Analysis in Apex Legends and MOBA Games

While direct research on character-specific performance in Apex Legends is limited, insights from other multiplayer online battle arena (MOBA) games provide valuable context. For example, da Costa Oliveira et al. [19] examined character selection and team composition in League of Legends, concluding that diverse team roles often enhance overall team performance. This principle applies to Apex Legends, where team synergy and optimal character selection contribute significantly to success.

Moreover, Lago [20] studied the impact of opponent quality and match status on strategy adaptation in football, a concept also relevant in battle royale games. In Apex Legends, players frequently shift between offensive and defensive strategies based on real-time game conditions, implying that adaptability may be a determinant of successful outcomes.

Finally, Dehpanah et al. [11] proposed a behavioral rating system for competitive shooters, incorporating features like aggressiveness, strategy, and teamwork to create a holistic understanding of performance. Given Apex Legends' emphasis on character-specific abilities and coordination, this approach offers a framework for evaluating player performance in team-based shooters.

C. Challenges and Gaps in Battle Royale Performance Metrics

Traditional metrics such as kill-death ratio (K/D) or win rates often fall short in capturing the complexities of performance in battle royale games. As Dehpanah et al. [7] found, these metrics lack the nuance to measure inter-player dependencies and fast-paced, team-centric dynamics integral to success in games like Apex Legends.

These limitations underscore the need for a refined approach to performance evaluation in Apex Legends. This study aims to address these gaps by incorporating character-specific metrics that better reflect individual contributions to team success.

By focusing on metrics such as kills per win and kills per match within different Legend categories, this research seeks to provide a more comprehensive understanding of player performance in this genre.

D. Clustering and Statistical Analysis in Gaming Research

Clustering has been widely used in gaming research to identify player types, optimise matchmaking, and enhance game balance. Bauckhage et al. [14] demonstrated how clustering can reveal hidden patterns in player data, such as preferred strategies or performance tiers, allowing developers and researchers to better understand player engagement and behavior. Sifa et al. [15] expanded on this by analysing player behavior across multiple games on Steam, showcasing clustering's ability to uncover behavioral archetypes applicable across different game genres.

Additionally, statistical analysis methods are crucial for quantifying performance correlations and testing hypotheses, which provides insights into factors influencing player success. For instance, Guzdial et al. [17] used statistical methods to predict player outcomes, while Drachen et al. [21] analysed gameplay metrics to understand user engagement and behavior patterns in gaming.

These techniques align well with the goals of this research, where clustering is applied to categorise Legends by performance trends, and statistical analysis is used to examine relationships between variables like kills and wins. Such methods provide a robust framework for studying the complex dynamics of team-based games.

E. Regression and Correlation in Player Performance Analysis

Regression and correlation analysis are widely used in gaming research to examine relationships between in-game metrics and player success, providing insights into factors that contribute to performance outcomes. For instance, Pearson correlation has been used to measure the strength of associations between variables like kills, deaths, and assists, helping to identify performance indicators in competitive games [22]. Linear regression models are also applied to predict outcomes based on multiple in-game variables, offering a structured approach to quantify how certain metrics, such as offensive actions or teamwork, impact win rates [23].

In this study, correlation analysis helps evaluate the relationship between offensive metrics and player success, while regression is employed to predict success outcomes based on various performance indicators.

V. METHODOLOGY

This section outlines the data collection, preprocessing, clustering, and statistical techniques used to investigate player performance in *Apex Legends*. We aim to evaluate the impact of Legend categories on player success metrics and examine relationships between offensive performance and win metrics across player career data.

A. Ethical Considerations

The player statistics used in this study were sourced from publicly accessible leaderboards on the *Apex Legends Status* website [24]. As this data is open to the public, it was appropriate for analysis in our study. To maintain player privacy, no gamer tags, usernames, or identifiable information were included in this report, ensuring that individual players remain anonymous within our analysis.

B. Analysis Workflow

The analysis pipeline, as shown in Figure 1, was split into two branches to address the hypotheses:

- **Hypothesis 1:** Clustering techniques were applied to group Legends by performance, followed by statistical tests to evaluate differences in metrics like Kills per Match and Kills per Win.
- **Hypothesis 2:** Correlation analysis assessed relationships between Career Kills and Wins, while regression analysis modeled the predictive relationship between these metrics.

C. Data Collection

Data was sourced from *Apex Legends Status*, a public API providing player statistics [24]. Different data collection strategies were implemented for each hypothesis to ensure relevance to specific performance metrics.

1) Hypothesis 1: Legend-Specific Data Collection: To analyse how Legend type influences performance, data on the top 500 players per Legend was collected via web scraping from public leaderboards [25], [24]. Collected statistics included:

- Legend Damage
- Legend Kills
- Legend Matches Played
- Legend Wins

Legends with incomplete win data (Alter, Ballistic, Newcastle, and Conduit) were omitted from win-related analyses to ensure data consistency.

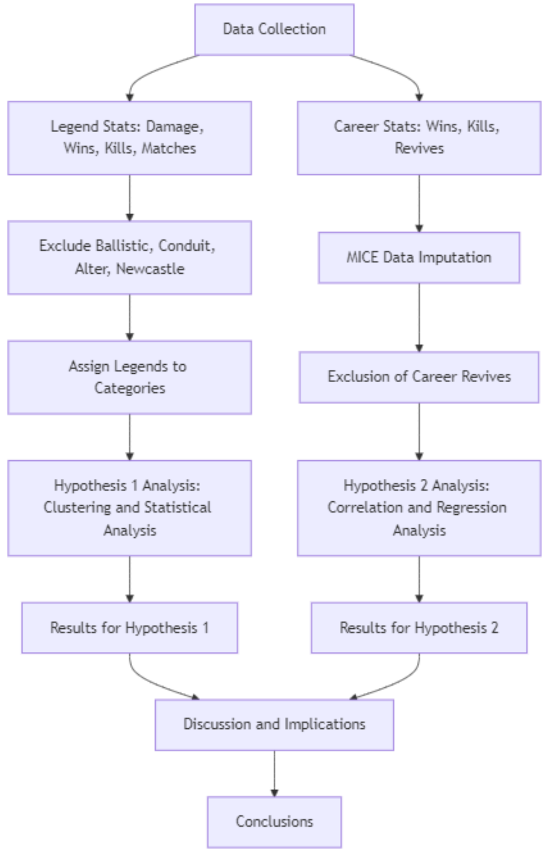

Fig. 1. Pipeline for analysing player performance in Apex Legends. The pipeline covers data collection, preprocessing, and hypothesis-specific analyses using clustering, statistical tests, correlation, and regression methods.

2) Hypothesis 2: Career Data Collection: For Hypothesis 2, which examines the relationship between Career Kills and Career Wins, data on the top 900 players per platform (Playstation, PC, Xbox) was collected using unique identifiers (UIDs) via the API. The following career statistics were obtained:

- Career Kills
- Career Wins
- Career Revives

D. Data Preprocessing

1) Data Imputation for Hypothesis 2: For Hypothesis 2, which examines Career Kills and Career Wins, the API occasionally returned missing values in the career data. To handle these gaps, three imputation methods were tested to identify the most suitable approach:

- **Mean Imputation:** Missing values were replaced with column averages, though this approach may limit the preservation of variable relationships [26].
- **K-Nearest Neighbors (KNN) Imputation:** Missing values were filled using averages from the nearest data points based on similarity [27].
- **Multiple Imputation by Chained Equations (MICE):** A multivariate approach that preserves variable correlations using chained equations, often outperforming univariate imputation methods in complex datasets [28].

a) Evaluation of Imputation Methods: To select the most suitable imputation method for Hypothesis 2's career data, correlation heatmaps and scatter plots were generated for each method. Figures 2 and 3 illustrate the correlation structure and data relationships maintained with MICE, which was selected for preserving higher correlations between key metrics such as Career Kills, Career Wins, and Career Revives.

As shown in Figures 2 and 3, MICE demonstrated superior correlation preservation, supporting its selection for this study's career data analysis.

E. Legend Categorisation

Each Legend was grouped into one of five categories (III-E). Aggregated performance metrics were calculated for each category, allowing for comparative analysis across Legend types. This classification facilitated an analysis of category-specific performance trends.

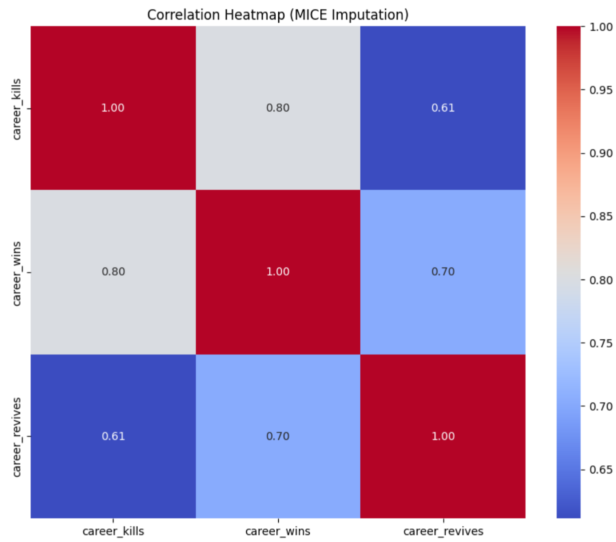

Fig. 2. Correlation Heatmap of Career Stats Using MICE Imputation

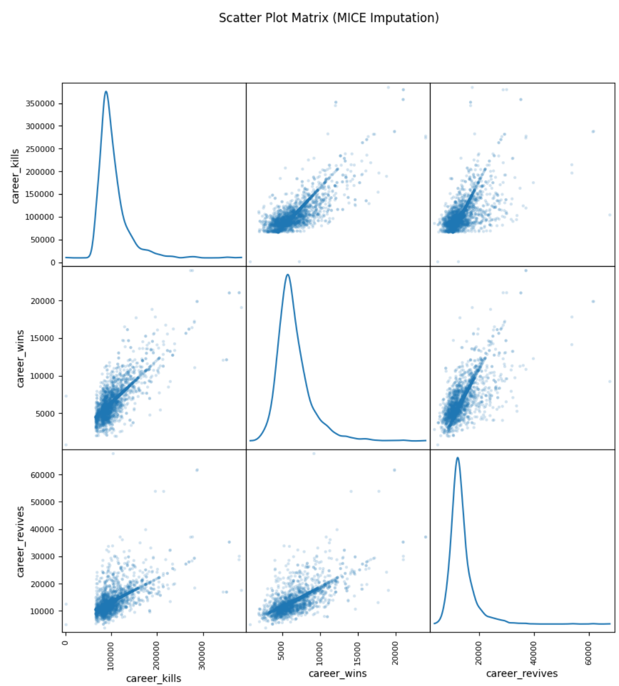

Fig. 3. Scatter Plot Matrix of Career Stats Using MICE Imputation

F. Performance Metrics

Each hypothesis used specific performance metrics aligned with its objectives.

1) Metrics for Hypothesis 1: To explore Legend category performance in Hypothesis 1, two key metrics were selected:

- **Kills per Match (*Kills_per_Match*):** Reflects offensive capabilities, showing average kills per match. Calculated as:

$$Kills\_per\_Match = \frac{\text{Total Kills}}{\text{Total Matches Played}} \quad (1)$$

- **Kills per Win (*Kills_per_Win*):** Measures kill efficiency related to wins, calculated as:

$$Kills\_per\_Win = \frac{\text{Total Kills}}{\text{Total Wins}} \quad (2)$$

a) Exclusion of Damage per Match: *Damage_per_Match* was excluded from Hypothesis 1 due to its high correlation with kill metrics and less direct alignment with study objectives, ensuring a focused analysis on kill-based performance.

2) Metrics for Hypothesis 2: For Hypothesis 2, two main metrics were chosen to examine correlations between offensive performance and success:

- **Career Kills:** A cumulative measure of offensive performance across matches, representing the player's ability to secure kills.
- **Career Wins:** Total wins, used to determine its relationship with Career Kills as an indicator of success.

a) Exclusion of Career Revives: Although *Career Revives* was initially collected, it was excluded from the primary analysis for Hypothesis 2. This decision was made to maintain a focused examination of offensive metrics (Career Kills) and success outcomes (Career Wins). *Career Revives* primarily reflects supportive play rather than offensive performance, and its inclusion could introduce confounding effects not directly aligned with the objectives of Hypothesis 2.

G. Clustering Analysis for Hypothesis 1

Clustering was applied as an unsupervised learning method to identify performance patterns across Legend categories [29]. In unsupervised learning, clustering organises data points into groups (clusters) based on similarity without predefined labels, allowing us to uncover underlying patterns in the dataset [30], [14].

Clustering helps to group Legends with similar performance characteristics, which may reveal category-specific playstyles or performance trends. By analysing these clusters, we can assess whether Legends within the same category (III-E) exhibit similar gameplay behaviors, providing insights into the impact of Legend roles on player performance. This unsupervised approach is suitable as it enables pattern discovery without prior assumptions about how Legends should group together.

1) Data Preparation for Clustering:

- **Data Cleaning and Preprocessing:** Legends with incomplete metrics were removed to ensure data consistency

(V-C1). Outliers were identified and excluded based on Z-scores, calculated as:

$$Z = \frac{X - \mu}{\sigma} \tag{3}$$

where X is the data point, μ is the mean, and σ is the standard deviation. Data points with $|Z| > 3$ were excluded as outliers, as they may distort clustering results by introducing extreme values [31].

- **Feature Scaling:** To ensure that Kills per Match and Kills per Win contributed equally to the clustering process, these features were standardised using StandardScaler. This transformation centered each feature to have a mean of 0 and a standard deviation of 1, calculated as:

$$X_{scaled} = \frac{X - \mu}{\sigma} \tag{4}$$

where X represents the original feature value, μ is the mean of the feature, and σ is the standard deviation. Standardising the data in this way prevents any single feature from disproportionately influencing the clustering algorithm due to differing scales [32].

- **Dimensionality Reduction with PCA:** Principal Component Analysis (PCA) was applied to reduce the dimensionality of the dataset, facilitating clear visualisation of clusters in a two-dimensional space. PCA transforms the data into a new coordinate system where the axes (principal components) capture the maximum variance, enabling us to simplify the data while preserving key patterns [33].
PCA achieves this transformation by calculating the eigenvalues and eigenvectors of the data covariance matrix. The data is then projected onto the principal components, as expressed by:

$$Z = X \cdot W \tag{5}$$

where Z represents the transformed data, X is the original dataset, and W contains the eigenvectors of the covariance matrix.
In this study, PCA reduced the dataset to two principal components. This reduction allowed us to visualise the clustering results in 2D space, enhancing interpretability by illustrating cluster separations without directly influencing the clustering process itself.

2) Optimal Number of Clusters: To determine the optimal cluster count (K), two methods were used, with results visualised in Figures 4 and 5.

- **Elbow Method:** This method involves plotting the Within-Cluster Sum of Squares (WCSS) across different K values. The "elbow" point, where the rate of decrease sharply slows, indicates an ideal balance between cluster compactness and simplicity. As shown in Figure 4, an elbow was identified at $K = 3$, suggesting three clusters as optimal [34].
- **Silhouette Analysis:** This measures how similar each point is to its own cluster compared to other clusters, providing an indication of cluster cohesion. Scores that are high imply well-separated and cohesive clusters. Figure 5 shows that silhouette scores are maximised at $K = 3$, supporting this choice for the optimal cluster count [35].

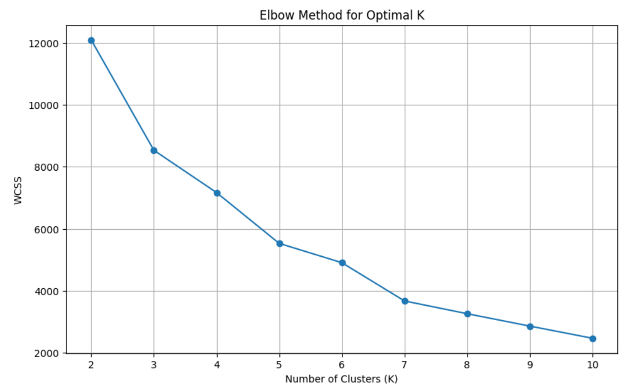

Fig. 4. Elbow Method Plot for Determining Optimal K

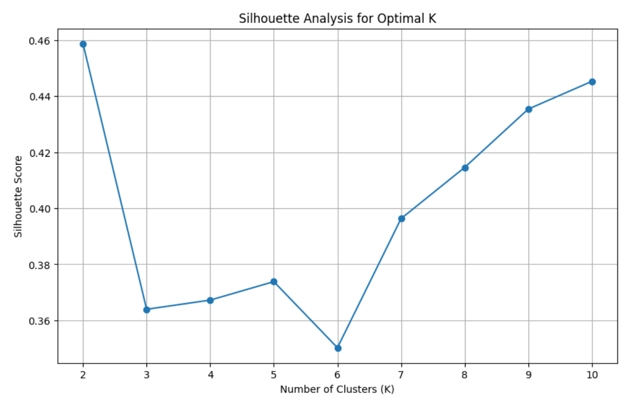

Fig. 5. Silhouette Scores Across K Values

3) Clustering Algorithms: Three clustering algorithms were applied to identify performance patterns across Legend categories. Each method offers unique benefits that aid in exploring different clustering perspectives.

- **K-Means Clustering:** A partition-based algorithm that divides data into K clusters by minimising within-cluster variance. It iteratively updates cluster centroids by minimising the sum of squared distances between each data point x_i and its assigned cluster center μ_k:

$$\sum_{k=1}^{K} \sum_{x_i \in C_k} ||x_i - \mu_k||^2 \tag{6}$$

where C_k represents the k-th cluster, and μ_k is its centroid. The "k-means++" initialisation was used to place initial centroids strategically, improving convergence speed and cluster quality [36].

- **Gaussian Mixture Models (GMM):** A probabilistic approach that assumes each cluster is generated from a Gaussian distribution [37]. GMM calculates the likelihood that a data point belongs to each cluster, providing soft clustering. The likelihood for data point x belonging to cluster k is given by:

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k) \tag{7}$$

where π_k is the weight (or prior) of the k-th Gaussian, μ_k is the mean, and Σ_k is the covariance matrix for each cluster [38].
- **Hierarchical Clustering:** An agglomerative method that builds a nested hierarchy of clusters. Using Ward linkage, it minimises the within-cluster variance at each step [39]. For two clusters A and B, the distance $D(A, B)$ is defined to minimise the increase in total within-cluster variance:

$$D(A, B) = \frac{|A| \cdot |B|}{|A| + |B|} ||\mu_A - \mu_B||^2 \tag{8}$$

where $|A|$ and $|B|$ are the number of points in clusters A and B, and μ_A and μ_B are their centroids. This technique is useful for visualising relationships between clusters through a dendrogram [40].

4) Evaluation of Clustering Results: To assess cluster quality, both internal and external metrics were applied, each providing insights into different aspects of clustering performance:

- **Silhouette Score:** An internal metric that evaluates how similar each data point is to its own cluster compared to other clusters. Higher scores indicate well-defined clusters with greater cohesion and separation. The Silhouette Score $S(i)$ for each point i is calculated as:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \tag{9}$$

where $a(i)$ is the average distance between i and other points in the same cluster, and $b(i)$ is the minimum average distance between i and points in a different cluster [35].
- **Calinski-Harabasz Index:** Also known as the Variance Ratio Criterion, this index measures the ratio of between-cluster dispersion to within-cluster dispersion. Higher values suggest distinct and compact clusters. The index CH is calculated as:

$$CH = \frac{\text{trace}(B_k)/(k-1)}{\text{trace}(W_k)/(n-k)} \tag{10}$$

where B_k is the between-cluster scatter matrix, W_k is the within-cluster scatter matrix, k is the number of clusters, and n is the total number of points [41].
- **Adjusted Rand Index (ARI):** An external metric that compares the clustering results to predefined categories (in this case, Legend categories) to measure alignment with known labels. The ARI is calculated as:

$$ARI = \frac{\sum_{ij} \binom{n_{ij}}{2} - \left[\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}\right] / \binom{n}{2}}{\frac{1}{2}\left[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}\right] - \left[\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}\right] / \binom{n}{2}} \tag{11}$$

where n_{ij} is the number of elements in the intersection between the i-th and j-th clusters, a_i and b_j are the sums of elements for the i-th row and j-th column respectively, and n is the total number of points [42].

H. Statistical Analysis for Hypothesis 1

The statistical analysis aimed to evaluate differences in performance metrics across Legend categories (III-E). Since the data showed a non-normal distribution [43], non-parametric [44] tests were employed to ensure robustness:

- **Kruskal-Wallis H Test:** A non-parametric alternative to ANOVA, this test assesses differences in medians across multiple independent groups. The Kruskal-Wallis statistic H is calculated as:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{g} n_i \left(R_i - \frac{N+1}{2}\right)^2 \tag{12}$$

where N is the total number of observations, g is the number of groups, n_i is the number of observations in group i, and R_i is the rank sum for group i. This test was applied to both Kills per Match and Kills per Win to determine if significant differences exist among Legend categories [45].
- **Dunn's Post-hoc Test:** Following a significant Kruskal-Wallis result, Dunn's test with Bonferroni correction was used for pairwise comparisons to identify specific Legend categories that differ significantly. The test statistic Z_{ij} for comparing groups i and j is given by:

$$Z_{ij} = \frac{R_i - R_j}{\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \tag{13}$$

where R_i and R_j are the average ranks for groups i and j, respectively, and N is the total sample size [46].
- **Cliff's Delta:** As a measure of effect size, Cliff's Delta (δ) was calculated to assess the practical significance of observed differences. This non-parametric measure is defined as:

$$\delta = \frac{\text{\# of pairs where } X > Y - \text{\# of pairs where } X < Y}{n_X \cdot n_Y} \tag{14}$$

where n_X and n_Y are the sample sizes of the two groups being compared. Cliff's Delta provides a sense of how often values in one group are greater than those in another, offering additional context beyond p-values [47].

I. Statistical Analysis for Hypothesis 2

For Hypothesis 2, various statistical methods were applied to analyse the relationship between Career Kills and Career Wins, focusing on both linear and non-linear associations.

1) Correlation Analysis: To assess the strength and direction of the relationship between Career Kills and Career Wins, three correlation measures were employed:

- **Pearson Correlation:** This parametric test was used to evaluate the linear relationship between Career Kills and Career Wins. The Pearson correlation coefficient r ranges from -1 to 1, where values closer to 1 or -1 indicate a stronger linear relationship. It is calculated as:

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \tag{15}$$

  where X and Y represent Career Kills and Career Wins, respectively [48].
- **Spearman's Rank Correlation:** A non-parametric measure that evaluates the monotonic relationship between Career Kills and Career Wins by ranking data points. Spearman's rank correlation ρ is calculated based on the ranks of X and Y, making it more robust against outliers and non-normal data distributions. It is computed as:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{16}$$

  where d_i is the difference between ranks of X and Y, and n is the number of observations [49].
- **Kendall's Tau:** Another non-parametric measure that assesses the strength of the relationship between two variables by comparing the concordance and discordance of data pairs. Kendall's Tau τ is calculated as:

$$\tau = \frac{C - D}{\sqrt{(C + D + T) \cdot (C + D + U)}} \tag{17}$$

  where C is the number of concordant pairs, D is the number of discordant pairs, T is the number of ties in X, and U is the number of ties in Y [49], [50].

2) Regression Analysis: To model the relationship between Career Wins and Career Kills, a simple linear regression model was applied. This approach was selected due to the straightforward nature of the relationship observed between these two metrics, as suggested by the correlation analysis.

- **Simple Linear Regression:** This model was chosen to assess the direct, linear relationship between Career Kills and Career Wins. It is defined as:

$$\text{Career Wins} = \beta_0 + \beta_1 \cdot \text{Career Kills} + \epsilon \tag{18}$$

  where β_0 is the intercept, β_1 is the coefficient representing the effect of Career Kills on Career Wins, and ϵ represents the error term [51].

Simple linear regression was deemed appropriate as the correlation analysis revealed a strong positive association between Career Kills and Career Wins, justifying the use of a model that captures this linear relationship effectively. This approach also facilitates interpretability, providing clear insights into the predictive power of Career Kills on Career Wins without additional model complexity.

3) Assumption Checks: To ensure the reliability and validity of the regression results, assumption checks were conducted for key aspects of the model:

- **Normality (Shapiro-Wilk Test):** This test assesses whether the residuals of the regression model follow a normal distribution. The Shapiro-Wilk test statistic W is calculated as:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{19}$$

  where $x_{(i)}$ represents the ordered residuals, a_i are constants generated from the means, variances, and covariances of the sample, and $\bar{x}$ is the sample mean. A low W value and corresponding p-value suggest a departure from normality [52].
- **Homoscedasticity (Breusch-Pagan Test):** The Breusch-Pagan test checks for constant variance in residuals across predictor levels. It regresses the squared residuals on the predictor(s) to test for non-constant variance. The test statistic χ^2 is calculated as:

$$\chi^2 = \frac{n \cdot R^2}{2} \tag{20}$$

  where n is the sample size and R^2 is the coefficient of determination from the auxiliary regression of squared residuals. A significant χ^2 value indicates heteroscedasticity [53].

J. Conclusion

In summary, the methodological approach combines data collection, preprocessing, clustering, and regression analysis to systematically investigate player performance patterns in *Apex Legends*. Clustering algorithms and statistical tests were carefully chosen to explore performance differences across Legend categories, while robust regression models were applied to examine relationships between Career Kills and Career Wins. Assumption checks were conducted to ensure the validity and reliability of the results. This methodological framework provides a comprehensive foundation for analysing performance metrics and deriving actionable insights in the subsequent sections.

VI. Results

1) Cluster Evaluation Metrics: Table I presents the evaluation metrics for each clustering method used in this analysis. The metrics include the Silhouette Score (9), Calinski-Harabasz Index (10), and Adjusted Rand Index (ARI) (11), as outlined in the Methodology section.

Among the clustering methods applied—K-Means (6), Gaussian Mixture Models (7), and Hierarchical Clustering (8)—**Hierarchical Clustering** achieved the highest Adjusted Rand Index of 0.0918, indicating the best alignment with the known Legend roles. It also demonstrated strong internal cohesion, as shown by its competitive Silhouette Score (9) and Calinski-Harabasz Index (10). While K-Means produced a marginally higher Calinski-Harabasz score, Hierarchical Clustering's performance across multiple metrics suggests a more well-defined cluster structure and meaningful separation of Legends.

TABLE I
CLUSTERING EVALUATION METRICS FOR LEGEND CATEGORIES

Clustering Method	Silhouette Score	Calinski-Harabasz Index	Adjusted Rand Index (ARI)
K-Means	0.3639	7956.83	0.0654
GMM	0.2694	4753.73	0.0396
Hierarchical	**0.3396**	**7071.14**	**0.0918**

a) Conclusion: Based on these evaluations, **Hierarchical Clustering** was selected as the most suitable method for this analysis due to its optimal combination of internal cohesion and alignment with predefined Legend categories. This choice provides a robust foundation for interpreting performance patterns within each Legend category in subsequent analyses.

2) Cluster Visualisation with PCA: To visualise the clusters, Principal Component Analysis (PCA) (5) was used to reduce the dimensionality of the dataset to two principal components. This approach enables an interpretable view of the clustering results, where each data point represents a Legend's performance profile based on key metrics, specifically Kills per Match and Kills per Win. Figure 6 displays the PCA-reduced visualisation of clusters formed using Hierarchical Clustering.

The PCA visualisation reveals three distinct clusters, suggesting natural groupings of Legends based on performance metrics. The separation between clusters indicates that Hierarchical Clustering effectively differentiates Legends into groups with unique performance characteristics, aligning with the predefined Legend categories. This clustering outcome suggests meaningful variations in playstyle and success metrics across different Legend types.

3) Cluster Profiles: Table II summarises the average values for key performance metrics within the three identified clusters. These profiles highlight significant differences in Legend performance across clusters.

a) Interpretation of Cluster Profiles: The analysis of the cluster profiles reveals distinct performance characteristics:

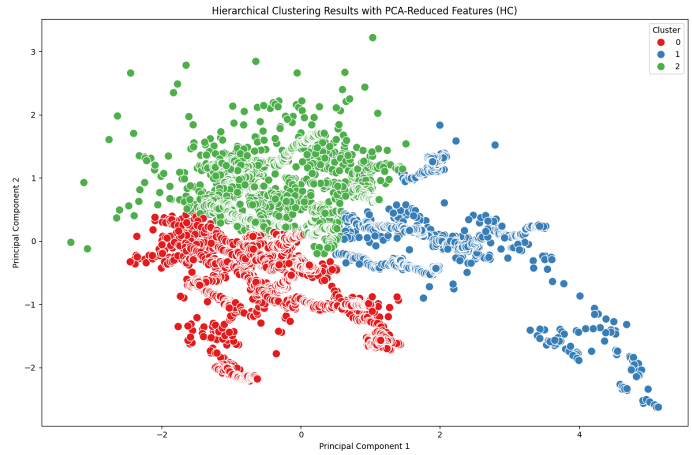

Fig. 6. Hierarchical Clustering results visualised with PCA-reduced features. Each color represents a different cluster, illustrating the separation and cohesion of Legends' performance profiles.

TABLE II
HIERARCHICAL CLUSTERING CLUSTER PROFILES

Cluster	Mean Kills per Match	Mean Kills per Win
0	4.28	19.52
1	6.24	22.81
2	5.22	16.90

- **Cluster 0:** Legends in this cluster exhibit lower kills per match but higher kills per win, suggesting an efficient playstyle that effectively converts kills into victories.
- **Cluster 1:** This cluster is characterised by Legends with the highest average kills per match and kills per win, indicating a strongly offensive-oriented playstyle.
- **Cluster 2:** Legends here demonstrate moderate values in both kills per match and kills per win, representing a balanced approach to performance metrics.

4) Legend Category Distribution within Clusters: To further investigate the alignment of clusters with the predefined Legend categories, we analysed the distribution of these categories within each cluster. Table III presents this distribution, while Figure 7 visually represents the results.

TABLE III
HIERARCHICAL CLUSTERING - LEGEND CATEGORY DISTRIBUTION PER CLUSTER (%)

Cluster	Assault (%)	Controller (%)	Recon (%)	Skirmisher (%)	Support (%)
0	10.99	26.80	29.97	21.85	10.38
1	49.57	8.93	16.69	24.81	0.00
2	15.65	6.23	8.00	35.86	34.26

a) Interpretation of Legend Distribution: The distribution of Legend categories within each cluster provides further insights into the performance characteristics:

- **Cluster 0:** Predominantly composed of Controller and

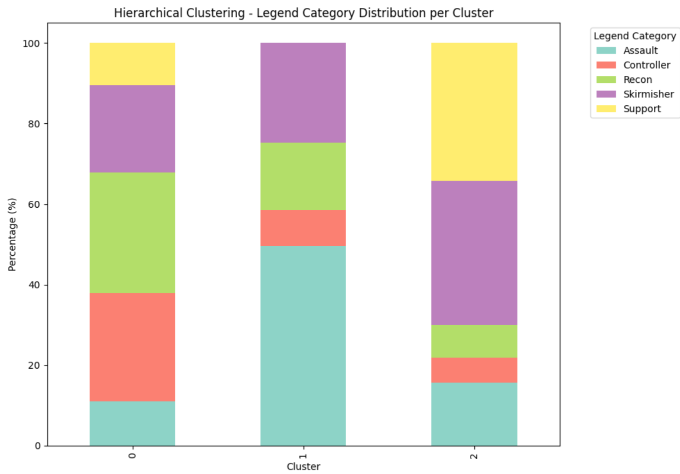

Fig. 7. Hierarchical Clustering - Legend Category Distribution per Cluster. Each color represents a different Legend category within the clusters.

Recon Legends, this cluster aligns with its efficient performance in terms of kills per win.

- **Cluster 1:** Characterised by a high concentration of Assault Legends, consistent with the offensive nature observed in the cluster profile.
- **Cluster 2:** Displays a balanced mix of Skirmisher and Support Legends, indicating utility-oriented gameplay with moderate kills per match.

A. *Hypothesis 1: Statistical Analysis*

This section presents the statistical analysis conducted to evaluate differences in Kills per Match and Kills per Win across Legend categories.

1) Kruskal-Wallis H Test: The Kruskal-Wallis H test (12) was conducted to assess whether significant differences exist across Legend categories for the two key metrics: Kills per Match and Kills per Win.

- **Kills per Match:** The Kruskal-Wallis test revealed significant differences among Legend categories, with $H = 568.53$ and $p < 0.001$, indicating substantial variation in Kills per Match across categories.
- **Kills per Win:** Similarly, the test for Kills per Win showed significant differences, with $H = 3098.52$ and $p < 0.001$, suggesting that Legend categories differ significantly in their efficiency of converting kills into wins.

The results of the Kruskal-Wallis test are visualised in Figure 8 and Figure 9, which display the distribution of Kills per Match and Kills per Win across Legend categories.

2) Dunn's Post-hoc Test and Cliff's Delta: To further explore pairwise differences among Legend categories, Dunn's post-hoc test with Bonferroni correction (13) was applied. Additionally, Cliff's Delta (14) was calculated as a measure of effect size to assess the practical significance of observed differences.

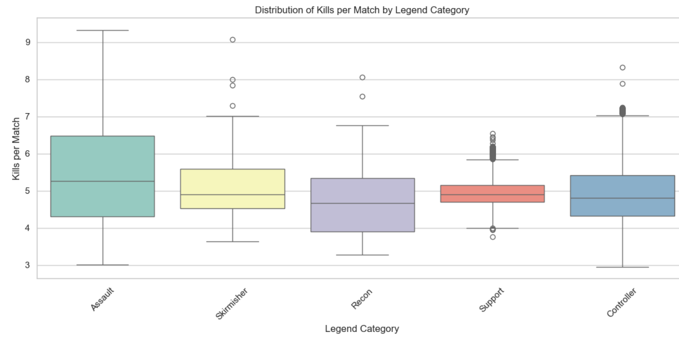

Fig. 8. Distribution of Kills per Match across Legend Categories

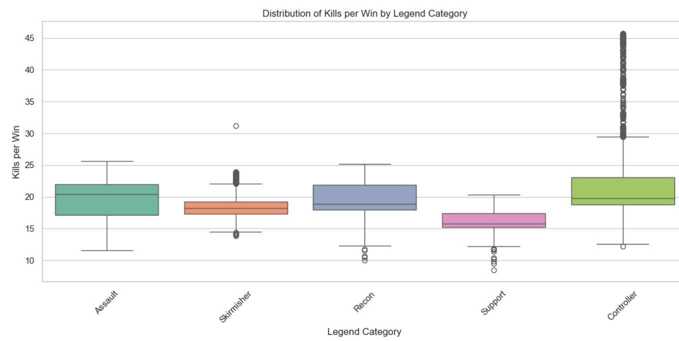

Fig. 9. Distribution of Kills per Win across Legend Categories

- **Kills per Match:** Significant differences were found between the Assault category and all other categories ($p < 0.05$), with Assault achieving the highest Kills per Match. Cliff's Delta indicated moderate to strong effect sizes between Assault and Recon (0.333), Assault and Support (0.325), and Assault and Skirmisher (0.279), suggesting meaningful practical differences. Recon also showed a moderate effect size over Support (-0.265), supporting its higher performance relative to Support.
- **Kills per Win:** All pairwise comparisons yielded significant differences ($p < 0.05$), with Assault leading, followed by Controller, Recon, Skirmisher, and Support. Cliff's Delta values indicated strong effects, particularly between Support and Controller (-0.786), Assault and Support (0.753), and Recon and Support (0.681). The large effect sizes confirm that these categories differ considerably in their kill-to-win efficiency.

In summary, the statistical analysis reveals a significant distinction in offensive performance across Legend categories, with Assault Legends showing the highest values in both Kills per Match and Kills per Win. These findings, supported by statistical significance and strong effect sizes, highlight a distinct performance advantage for Assault Legends, suggesting their effectiveness in offensive roles.

B. Hypothesis 2: Career Kills and Wins Correlation

This section presents the correlation (V-I1) and regression (V-I2) analyses conducted to explore the relationship between Career Kills and Career Wins, including relevant assumption checks to validate the reliability of the results.

1) Correlation Analysis: The correlation between Career Kills and Career Wins was assessed using Pearson (15), Spearman's Rank (16), and Kendall's Tau (17) methods, all indicating a significant positive relationship that supports Hypothesis 2.

- **Pearson Correlation:** Career Kills and Career Wins demonstrated a strong positive linear correlation ($r = 0.7997, p < 0.001$), suggesting that higher kill counts are strongly associated with increased wins. This relationship is visualised in the correlation heatmap in Figure 10, illustrating a robust association between these two metrics.
- **Spearman's Rank Correlation:** The Spearman correlation, which assesses monotonic relationships, also showed a strong positive association between Career Kills and Career Wins ($\rho = 0.7440, p < 0.001$), further supporting a consistent trend across these variables.
- **Kendall's Tau Correlation:** Kendall's Tau provided additional confirmation of the positive relationship, with a correlation of $\tau = 0.5889, p < 0.001$. This result reflects the strength of the association even when accounting for rank-based non-parametric measures.

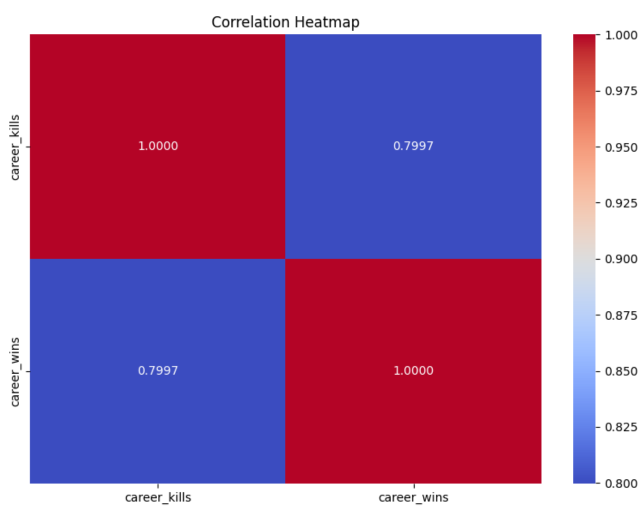

Fig. 10. Correlation heatmap displaying the relationship between Career Kills and Career Wins.

Together, these correlation results provide strong evidence of a positive relationship between Career Kills and Career Wins, demonstrating consistency across both parametric and non-parametric methods.

2) Regression Analysis: To further examine the relationship between Career Kills and Career Wins, a regression analysis was conducted.

a) Simple Linear Regression: A simple linear regression model (18) was fitted with Career Wins as the dependent variable and Career Kills as the predictor. This model explained 63.9% of the variability in Career Wins ($R^2 = 0.639$), indicating a strong linear relationship between the two metrics. The scatter plot in Figure 11 includes the regression line, confirming the predictive strength of Career Kills for Career Wins.

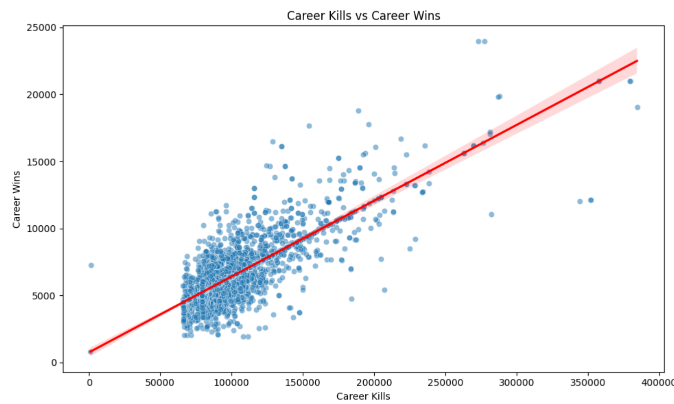

Fig. 11. Scatter plot with regression line illustrating the relationship between Career Kills and Career Wins.

b) Actual vs Predicted Career Wins: To assess the model's predictive accuracy, the Actual vs Predicted scatter plot (Figure 12) was generated. The alignment of points along the diagonal line suggests that the model provides a reliable prediction of Career Wins based on Career Kills.

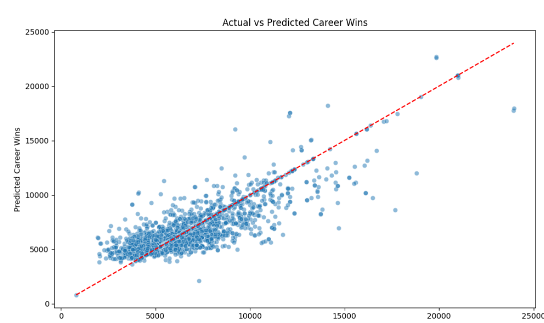

Fig. 12. Scatter plot of Actual vs Predicted Career Wins, showing a strong alignment along the diagonal, which indicates high predictive accuracy of the model.

3) Assumption Verification: Assumption checks were performed to validate the regression model's reliability and ensure robust interpretation of results:

- **Normality of Residuals (19):** The Shapiro-Wilk test ($p < 0.001$) suggested deviations from normality, further supported by the Q-Q plot in Figure 13, which shows significant departures from normality. These deviations imply some non-normality in the error distribution.
- **Homoscedasticity (20):** The Breusch-Pagan test ($p < 0.001$) indicated heteroscedasticity, which is visually supported by the residual plot in Figure 14. The plot shows increasing variability at higher kill counts, suggesting that prediction accuracy may vary across the range of Career Kills.

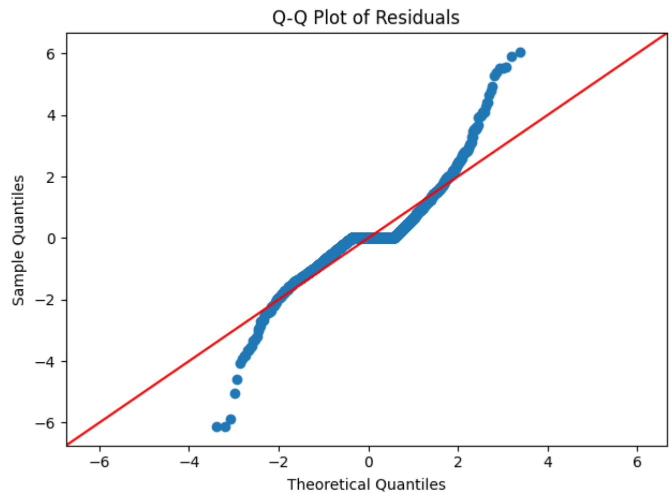

Fig. 13. Q-Q plot of residuals indicating non-normality in the regression residuals.

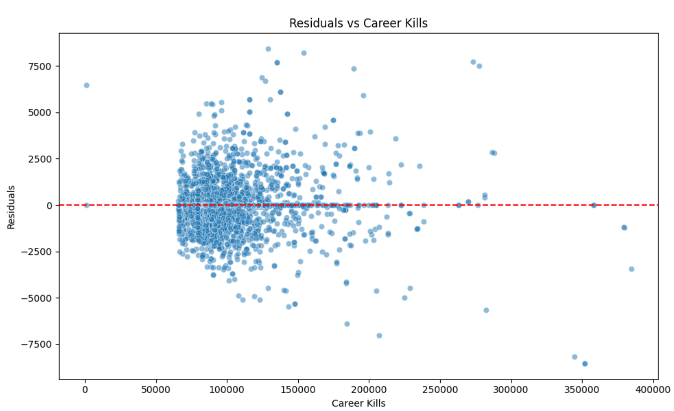

Fig. 14. Residual plot showing heteroscedasticity in the residuals across different levels of Career Kills.

In summary, the correlation and regression analyses support a strong positive relationship between Career Kills and Career Wins, as evidenced by both correlation coefficients and regression modeling. Although some assumption violations were noted, the overall predictive capability of Career Kills for Career Wins remains substantial, providing empirical support for Hypothesis 2.

VII. Discussion

A. Hypothesis 1: Clustering and Statistical Analysis

Hypothesis 1 proposed that different Legend categories would exhibit distinct performance profiles in terms of Kills per Match and Kills per Win, aligning with their predefined gameplay roles.

1) Cluster Profiles and Statistical Differences: The clustering analysis identified three clusters with unique performance characteristics:

- **Cluster 0** primarily comprises Controller and Recon Legends, characterised by lower kills per match but a high kills-to-win conversion rate. This profile aligns with a tactical role focused on efficiency rather than high kill counts.
- **Cluster 1**, predominantly composed of Assault Legends, demonstrates the highest kills per match and kills per win, reinforcing their offensive gameplay style and emphasis on aggressive engagements.
- **Cluster 2** includes Skirmisher and Support Legends, with moderate values for both metrics, indicative of a balanced approach prioritising versatility and utility.

The statistical analysis further supports these cluster profiles, with the Kruskal-Wallis test revealing significant differences across categories ($p < 0.001$) for both Kills per Match and Kills per Win. Dunn's post-hoc test indicated that Assault Legends significantly outperformed other categories in Kills per Match, while Support Legends ranked lowest, emphasising their focus on non-offensive contributions. These patterns align with the expected roles for each Legend category, suggesting that offensive Legends excel in kill metrics, whereas Support Legends focus on other gameplay objectives.

2) Alignment and Divergence between Analyses: Both clustering and statistical analyses provide consistent insights into the dominance of Assault Legends in offensive metrics. However, some nuances are observed:

- **Alignment**: The dominance of Assault Legends in Kills per Match and Kills per Win is evident across both analyses, affirming the alignment between expected roles and observed data.
- **Divergence**: The clustering analysis captures an overlap between Controller and Recon Legends within Cluster 0, suggesting a shared efficiency-focused gameplay style. The statistical analysis, however, reveals non-significant differences between Controller and Skirmisher in Kills per Match, hinting at versatility in these roles that clustering captures more distinctly.

This combination of alignment and divergence between clustering and statistical findings deepens the understanding of

Legend roles, confirming that while offensive Legends achieve higher kill metrics, other roles, such as Controller and Recon, maintain distinct efficiency-oriented profiles.

3) Integrated Interpretation and Hypothesis Validation: The findings from both clustering and statistical analyses provide strong support for Hypothesis 1. The alignment between clustering profiles and statistical significance tests validates the hypothesis that predefined gameplay roles influence performance metrics. Specifically, the offensive nature of Assault Legends is confirmed, as they consistently outperform other roles in kill-related metrics. The clustering analysis enriches this understanding by capturing nuanced groupings, such as the efficiency-oriented roles of Controller and Recon Legends.

In summary, Hypothesis 1 is validated by the combined insights from clustering and statistical results. Together, these analyses confirm that Legend categories align with specific performance profiles, particularly in terms of offensive and utility-focused gameplay styles.

B. Hypothesis 2: Career Kills and Wins Correlation

Hypothesis 2 posited a positive relationship between Career Kills and Career Wins, suggesting that offensive performance directly influences winning outcomes.

1) Correlation Analysis: The correlation analysis confirmed a strong, positive relationship between Career Kills and Career Wins, with a Pearson correlation coefficient of $r = 0.7997$ ($p < 0.001$). This finding supports Hypothesis 2, indicating that players with higher kill counts tend to achieve more wins. The correlation heatmap (Figure 10) visually reinforces this relationship, suggesting that offensive actions, such as securing kills, contribute significantly to winning outcomes.

2) Regression Analysis: The regression analysis provided additional evidence for Hypothesis 2. The simple linear regression model accounted for 63.9% of the variability in Career Wins ($R^2 = 0.639$), indicating that Career Kills is a strong predictor of Career Wins. Figure 11 illustrates this positive linear relationship, with the scatter plot and regression line confirming that players with higher kills are more likely to secure wins.

The ”Actual vs Predicted Career Wins” plot (Figure 12) shows close alignment along the diagonal, indicating that the model effectively predicts Career Wins based on Career Kills. This reinforces the hypothesis that offensive actions play a crucial role in determining success.

3) Assumption Verification and Limitations: While the results strongly support Hypothesis 2, assumption checks revealed some deviations. The Shapiro-Wilk test and Q-Q plot of residuals (Figure 13) indicated non-normality in the residuals, while the Breusch-Pagan test and residual plot (Figure 14) highlighted heteroscedasticity. These assumption violations suggest that while the relationship between Career Kills and Wins is robust, extreme values in offensive performance may introduce variability, potentially reducing model precision in predicting wins for outliers.

4) Integrated Interpretation and Hypothesis Validation: The consistency of findings from both correlation and regression analyses provides strong support for Hypothesis 2. The high correlation and predictive accuracy validate that Career Kills significantly correlate with Career Wins, underscoring the critical role of offensive performance in achieving winning outcomes.

In summary, the positive correlation between Career Kills and Career Wins strongly supports Hypothesis 2. The results confirm that kills are a key factor in competitive gameplay, affirming that offensive performance is integral to success.

C. Implications for Player Performance and Strategy

The findings from Hypothesis 1 demonstrate that distinct Legend categories exhibit specific performance profiles in terms of Kills per Match and Kills per Win, aligning closely with their intended gameplay roles. For instance, Assault Legends achieved the highest kill metrics across both clustering and statistical analyses, reinforcing their role as aggressive, offensive players. This suggests that players aiming for high-kill strategies or focused engagements may benefit most from selecting Assault Legends. Conversely, Controller and Recon Legends, which excelled in kills-to-win conversion rather than overall kills per match, may suit players prioritising tactical approaches and team-based support. Meanwhile, Support and Skirmisher Legends, which had the lowest metrics in both analyses, reinforce the notion of a more utility-focused role that contributes to team success beyond individual eliminations.

Hypothesis 2 further emphasises the importance of offensive actions for achieving career success, as seen in the positive correlation between Career Kills and Career Wins. This suggests that players who focus on improving their kill rates could potentially increase their overall wins, providing a clear objective for those seeking to enhance their performance. These results also offer insights for team compositions in competitive play, where a mix of offensive and tactical roles may balance high-kill engagements with strategic victories.

D. Limitations and Future Research Directions

While the analyses provided valuable insights, several limitations were encountered that could impact the generalisability and completeness of the findings. Additionally, there are numerous avenues for expanding upon this research to gain deeper insights into player performance and Legend effectiveness.

- **Data Limitations:** Apex Legends lacks an official API for accessing player statistics, necessitating the use of a third-party API. This limitation restricted available career stats and often returned missing values. Moreover, the analysis was constrained to the top 900 players across Playstation, PC, and Xbox due to the API's reliance on specific UIDs for player identification. Although this approach ensured data privacy, it limited the dataset to high-ranking players, potentially skewing results toward elite gameplay behaviors. Future research could address this limitation by incorporating data from other ranking brackets, examining whether observed patterns hold across different skill levels.
- **Limited Career Metrics:** The constraints of the third-party API allowed only Career Kills and Career Wins for analysis, excluding other relevant statistics such as assists, damage dealt, win streaks, or matches played. Expanding future studies to include a broader set of performance metrics could provide a more comprehensive understanding of player behavior and success factors. Additionally, exploring non-linear relationships, such as polynomial regressions, might offer insights into player progression and Legend effectiveness across skill levels.
- **Missing Data for Specific Legends:** Certain Legends, specifically Alter, Ballistic, Conduit, and Newcastle, lacked available win data, which necessitated their exclusion from parts of the analysis. This missing data may limit the generalisability of findings across all Legend roles and playstyles. Future research could focus on obtaining complete datasets that cover all Legends to ensure a more balanced evaluation of each role's effectiveness.
- **Web Scraping for Legend Data:** Data on the top 500 players per Legend was collected through web scraping from public leaderboards. Although this provided insights into Legend performance across categories, the scraping method could not guarantee data completeness or uniformity, limiting the dataset's granularity and robustness. Future studies could incorporate official data sources if available or use refined data collection methods to improve dataset consistency and validity.
- **Team Composition and Gameplay Synergies:** While this study focused primarily on individual Legend performance, team composition and coordinated gameplay modes could be significant factors in success. Future research could examine synergies between Legend roles in team settings to understand how combinations of roles enhance team success, particularly in competitive play.
- **Dynamic Nature of Apex Legends:** Apex Legends is continuously evolving, with regular updates introducing new Legends, weapons, and gameplay mechanics. These changes can significantly impact player strategies, Legend effectiveness, and overall gameplay balance. As a result, findings from this study may not fully capture the impact of future updates. Ongoing research that adapts to these changes could provide more timely insights, helping players and teams refine strategies in response to new content. Future studies could incorporate longitudinal analyses to track performance trends as new elements are added to the game.

VIII. Code Availability

The source code used for the analysis and results presented in this research is publicly available on GitHub. The repository can be accessed at the following link:

https://github.com/AltaafAlly/Apex-Legends-Research.

This repository includes scripts for data collection, preprocessing, clustering, statistical analysis, and visualisation of results.

IX. Conclusion

In this study, I provided a systematic analysis of *Apex Legends* player performance, focusing on the role-specific contributions of different Legend categories. This research offers several unique contributions to understanding player dynamics in team-based battle royale games. First, it is among the earliest to evaluate *Apex Legends* Legend roles through a combination of clustering and correlation analysis, moving beyond traditional metrics such as kill-death ratio to reveal deeper insights into role-specific performance.

My findings highlight that Assault Legends significantly outperform other categories in offensive metrics, specifically kills per match and kills per win. This pattern underscores the critical role of high-kill engagements in securing wins. The positive correlation observed between Career Kills and Career Wins reinforces the importance of offensive actions in achieving player success, a relationship not previously documented in this context.

Methodologically, this research demonstrates the effectiveness of combining clustering with statistical analysis to identify distinct performance trends across Legend roles, potentially offering a new approach for evaluating character roles in similar team-based games. My results provide actionable insights for players aiming to enhance performance through strategic Legend selection, as well as for developers looking to balance character abilities and refine matchmaking systems.

While this study focused on high-ranking players, future research could extend these methods to various ranking brackets to explore whether these performance trends hold across skill levels. This work establishes a foundation for continued exploration of gameplay dynamics in *Apex Legends*, offering a framework to examine how Legend roles influence player success.

X. Acknowledgments

I would like to express my heartfelt gratitude to my supervisors, Dr. Branden Ingram and Dr. Pravesh Ranchod, for their invaluable guidance, encouragement, and support throughout this research. Their expertise and insights were instrumental in shaping the direction of my study and ensuring its success. I am deeply appreciative of their mentorship, which has been both motivating and enlightening.

I also extend my thanks to the team at *Apex Legends Status* for providing open access to player statistics. This access was crucial for the data analysis in this project and enabled a deeper exploration of player performance patterns.

References

[1] A. Visti, T. N. Joelsson, and J. Smed, "Beyond skill-based rating systems: Analyzing and evaluating player performance," *Proceedings of the 21st International Academic Mindtrek Conference*, 2017.

[2] D. Carrilho, M. Santos Couceiro, J. Brito, P. Figueiredo, R. J. Lopes, and D. Araújo, "Using optical tracking system data to measure team synergic behavior: Synchronization of player-ball-goal angles in a football match," *Sensors*, vol. 20, no. 17, p. 4990, 2020.

[3] K. J. Shim, K. W. Hsu, S. Damania, C. DeLong, and J. Srivastava, "An exploratory study of player and team performance in multiplayer first-person-shooter games," *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, 2011.

[4] Y. Guo, S. Shen, O. Visser, and A. Iosup, "An analysis of online match-based games," *2012 IEEE International Workshop on Haptic Audio Visual Environments and Games (HAVE 2012)*, 2012.

[5] J. Fernandez-Navarro, L. Fradua, A. Zubillaga, and A. P. McRobert, "Evaluating the effectiveness of styles of play in elite soccer," *International Journal of Sports Science & Coaching*, vol. 14, no. 4, pp. 514–527, 2019.

[6] V. Joshy, "Openskill: A faster asymmetric multi-team, multiplayer rating system," *arXiv preprint arXiv:2401.05451*, 2024.

[7] A. Dehpanah, M. F. Ghori, J. Gemmell, and B. Mobasher, "The evaluation of rating systems in team-based battle royale games," *arXiv preprint arXiv:2105.14069*, 2021.

[8] M. Dick, O. Wellnitz, and L. Wolf, "Analysis of factors affecting players' performance and perception in multiplayer games," in *Proceedings of 4th ACM SIGCOMM workshop on Network and system support for games*, 2005, pp. 1–7.

[9] A. L. Wiki, "Legend," 2024, accessed: Oct. 21, 2024. [Online]. Available: https://apexlegends.fandom.com/wiki/Legend

[10] A. Becker and D. Görlich, "What is game balancing?-an examination of concepts," *ParadigmPlus*, vol. 1, no. 1, pp. 22–41, 2020.

[11] A. Dehpanah, M. F. Ghori, J. Gemmell, and B. Mobasher, "Behavioral player rating in competitive online shooter games," *arXiv preprint arXiv:2207.00528*, 2022.

[12] S. Ohno, "The link between battle royale games and aggressive feelings, addiction, and sense of underachievement: Exploring esports-related genres," *International Journal of Mental Health and Addiction*, vol. 20, no. 3, pp. 1873–1881, 2022.

[13] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[14] C. Bauckhage, A. Drachen, and R. Sifa, "Clustering game behavior data," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 7, no. 3, pp. 266–278, 2014.

[15] R. Sifa, A. Drachen, and C. Bauckhage, "Large-scale cross-game player behavior analysis on steam," in *Proceedings of the 11th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2015, pp. 198–204.

[16] A. Drachen, R. Sifa, C. Bauckhage, and C. Thurau, "Guns, swords and data: Clustering of player behavior in computer games in the wild," in *Proceedings of IEEE Conference on Computational Intelligence and Games*, 2012, pp. 163–170.

[17] M. Guzdial, B. Li, S. Lee, and M. O. Riedl, "Predicting player goals in open-ended digital games with markov chains," in *Proceedings of the 12th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2016, pp. 111–117.

[18] Q. He, D. Araújo, K. Davids, Y. H. Kee, and J. Komar, "Functional adaptability in playing style: A key determinant of competitive football performance," *Adaptive Behavior*, vol. 31, no. 6, pp. 545–558, 2023.

[19] V. da Costa Oliveira, B. J. Placides, M. d. F. O. Baffa, and A. F. da Veiga Machado, "A hybrid approach to build automatic team composition in league of legends," *Proceedings of SBGames*, 2017.

[20] C. Lago, "The influence of match location, quality of opposition, and match status on possession strategies in professional association football," *Journal of sports sciences*, vol. 27, no. 13, pp. 1463–1469, 2009.

[21] A. Drachen and A. Canossa, "Towards gameplay analysis via gameplay metrics," in *Proceedings of the 13th international MindTrek conference: Everyday life in the ubiquitous era*, 2009, pp. 202–209.

[22] A. R. Novak, K. J. Bennett, M. A. Pluss, and J. Fransen, "Performance analysis in esports: modelling performance at the 2018 league of legends world championship," *International Journal of Sports Science & Coaching*, vol. 15, no. 5-6, pp. 809–817, 2020.

[23] V. J. Hodge, S. Devlin, N. Sephton, F. Block, P. I. Cowling, and A. Drachen, "Win prediction in multiplayer esports: Live professional match prediction," *IEEE Transactions on Games*, vol. 13, no. 4, pp. 368–379, 2019.

[24] A. L. Status, "Apex legends status - stats and leaderboards," 2024, accessed: Oct. 21, 2024. [Online]. Available: https://apexlegendsstatus.com

[25] D. Glez-Peña, A. Lourenço, H. López-Fernández, M. Reboiro-Jato, and F. Fdez-Riverola, "Web scraping technologies in an api world," *Briefings in bioinformatics*, vol. 15, no. 5, pp. 788–797, 2014.

[26] A. R. T. Donders, G. J. Van Der Heijden, T. Stijnen, and K. G. Moons, "A gentle introduction to imputation of missing values," *Journal of clinical epidemiology*, vol. 59, no. 10, pp. 1087–1091, 2006.

[27] U. Pujianto, A. P. Wibawa, M. I. Akbar *et al.*, "K-nearest neighbor (k-nn) based missing data imputation," in *2019 5th International Conference on Science in Information Technology (ICSITech)*. IEEE, 2019, pp. 83–88.

[28] S. Van Buuren and K. Groothuis-Oudshoorn, "mice: Multivariate imputation by chained equations in r," *Journal of statistical software*, vol. 45, pp. 1–67, 2011.

[29] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[30] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, R. Tibshirani, and J. Friedman, "Unsupervised learning," *The elements of statistical learning: Data mining, inference, and prediction*, pp. 485–585, 2009.

[31] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, 2004.

[32] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*. Springer, 2013.

[33] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150202, 2016.

[34] M. A. Syakur, B. K. Khotimah, E. Rochman, and B. D. Satoto, "Integration k-means clustering method and elbow method for identification of the best customer profile cluster," in *IOP conference series: materials science and engineering*, vol. 336. IOP Publishing, 2018, p. 012017.

[35] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[36] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," Stanford, Tech. Rep., 2006.

[37] C. B. Do, "The multivariate gaussian distribution," *Section Notes, Lecture on Machine Learning, CS*, vol. 229, 2008.

[38] D. A. Reynolds *et al.*, "Gaussian mixture models." *Encyclopedia of biometrics*, vol. 741, no. 659-663, 2009.

[39] S. Sharma, N. Batra *et al.*, "Comparative study of single linkage, complete linkage, and ward method of agglomerative clustering," in

2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon). IEEE, 2019, pp. 568–573.
[40] F. Murtagh and P. Legendre, "Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion?" *Journal of classification*, vol. 31, pp. 274–295, 2014.
[41] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.
[42] L. Hubert and P. Arabie, "Comparing partitions," *Journal of classification*, vol. 2, pp. 193–218, 1985.
[43] K. L. Sainani, "Dealing with non-normal data," *Pm&r*, vol. 4, no. 12, pp. 1001–1005, 2012.
[44] S. Siegel, "Nonparametric statistics," *The American Statistician*, vol. 11, no. 3, pp. 13–19, 1957.
[45] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American statistical Association*, vol. 47, no. 260, pp. 583–621, 1952.
[46] A. C. Elliott and L. S. Hynan, "A sas® macro implementation of a multiple comparison post hoc test for a kruskal–wallis analysis," *Computer methods and programs in biomedicine*, vol. 102, no. 1, pp. 75–80, 2011.
[47] G. Macbeth, E. Razumiejczyk, and R. D. Ledesma, "Cliff's delta calculator: A non-parametric effect size program for two groups of observations," *Universitas Psychologica*, vol. 10, no. 2, pp. 545–555, 2011.
[48] K. Pearson, "Vii. note on regression and inheritance in the case of two parents," *proceedings of the royal society of London*, vol. 58, no. 347-352, pp. 240–242, 1895.
[49] G. E. Noether, "Why kendall tau?" *Teaching Statistics*, vol. 3, no. 2, pp. 41–43, 1981.
[50] R. H. A. Shiekh and E. F. El-Hashash, "A comparison of the pearson, spearman rank and kendall tau correlation coefficients using quantitative variables," *Asian Journal of Probability and Statistics*, pp. 36–48, 2022.
[51] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2021.
[52] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611, 1965.
[53] A. Abdul-Hameed and O. G. Matanmi, "A modified breusch–pagan test for detecting heteroskedasticity in the presence of outliers," *Retrieved online from the website: http://article. pamathj. net/pdf/10.11648. j. pamj*, vol. 20211006, 2021.