

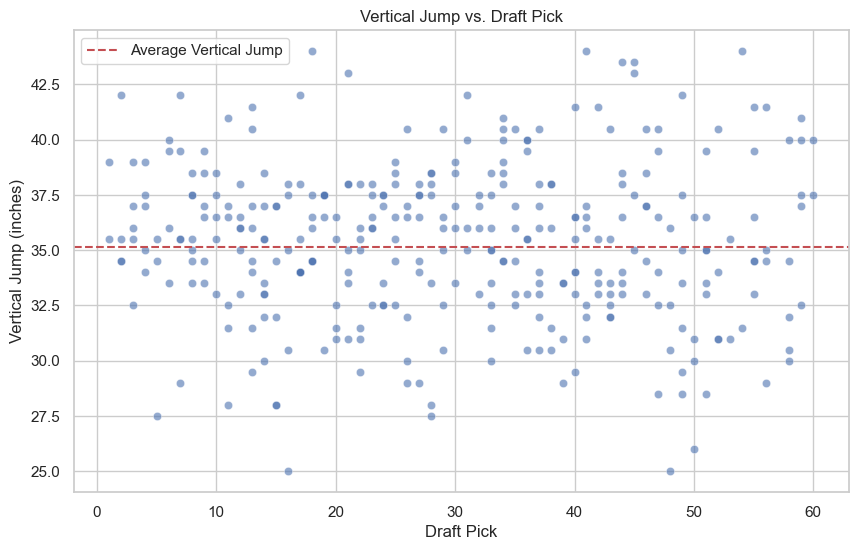
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | |
|  | | VerveBridge | | | | |  | |
|  |  | | | | | | |  |
|  | | | |  |  | | | |
|  | | | | Palak Singh |  | | | |
|  | | | | October 10, 2024—Data Analysis internship—Task 1 |  | | | |
|  | | |  | | |  | | |

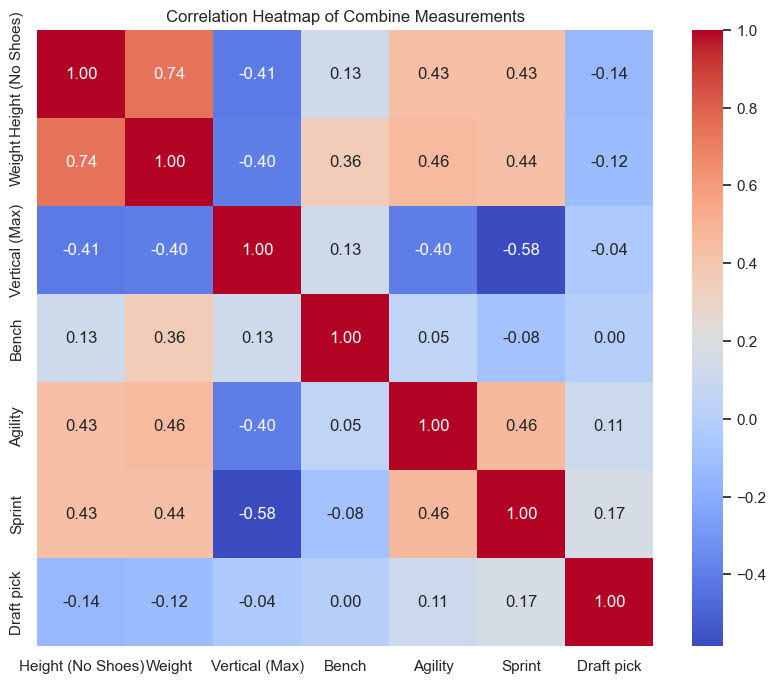
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | | |  |  | | |  |
|  | TASK 1 | | | | | | |  |
|  |  | | |  |  | | |  |
|  |  | |  | | |  | |  |
|  |  |  | Analysis on NBA Draft Combine Measurements  Objective:  The goal is to analyze NBA Draft Combine measurements to uncover patterns and insights that can help in predicting player success in the NBA. | | |  |  |  |
| Decorative | | | | |
|  |  |  |  |
|  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | |  |  | | |  | |  | |
|  | | | Research and Results | | | | | |  | |
|  | | |  |  | | |  | |  | |
|  | This report analyzes the NBA draft combine data to understand the relationship between player measurements and their success in the draft.  The following steps were undertaken for Analysis:   1. Data Collection: The dataset was obtained from a web-based platform and community for data scientists and machine learning practitioners. The dataset consisted of measurements from multiple NBA draft combines, including player height, weight, vertical jump, and other relevant metrics.   Here’s the source for dataset - https://www.kaggle.com/datasets/thedevastator/nba-draft-combine-measurement-data-from-2012-201?select=nba\_draft\_combine\_all\_years.csv  2. Data Cleaning: The dataset was cleaned by removing rows with missing values and filling empty cells with its mean value to ensure accurate analysis. Key features such as height, weight, vertical jump, and draft pick position were retained for analysis.  3. Descriptive Analysis: Key statistics (count, mean, standard deviation, minimum and maximum, Q1, Q2, Q3) were summarized to provide insights into the physical attributes of players. Trends and outliers were identified to understand variations among players.  4. Correlation Analysis: Correlations between different measurements and draft pick positions were examined using Pearson correlation coefficients to determine significant relationships.  5. Predictive Modeling: Linear regression models were developed to forecast player success based on their combine measurements.  6. Visualization: Various visualizations were created to present insights and patterns in the data effectively. | | | | | | | | |  |
|  | Findings-  Average Measurements:  - Height (No Shoes): Approximately 78.2 inches  - Weight: Around 217.5 pounds  - Vertical Jump (Max): About 33.7 inches  - Bench Press: Average of 10.5 repetitions  - Agility Drill Time: Approximately 11.1 seconds  Trends: There is a noticeable increase in vertical jump heights over the years, indicating a shift towards more athletic players.  Outliers: Notable outliers include Hasheem Thabeet (height of 85.25 inches) and DeJuan Blair (weight of 277 pounds), which may influence team evaluations.  Correlation Analysis   * Player Success (Draft Pick):   + Negative correlations (indicating better draft position is linked to higher performance):     - Vertical Jump (Max): Players with higher vertical jumps tend to be drafted higher (-0.35).     - Sprint Time: Faster sprint times show a mild correlation with a better draft pick (-0.17).   + Positive correlations (indicating worse draft position):     - Body Fat: Higher body fat percentages are weakly associated with lower draft picks (0.19), implying fitter players may be drafted earlier. * Player Height and Wingspan: There is a strong positive correlation between height and wingspan (0.87). Taller players generally have longer wingspans. * Weight and Vertical Jump: Heavier players tend to have lower vertical jumps, with a negative correlation of -0.58.   A linear regression model was developed using key features to predict draft pick positions, achieving an R² score of approximately 0.92, meaning about 92% of the variance in draft positions (a higher value indicates better performance). | | | | | | | | |  |
|  | Predictive Models Developed and Their Evaluation  In the analysis of the NBA Draft Combine data, a linear regression model was developed to predict player success as measured by their draft pick positions based on various combine measurements. Below are the details of the predictive models developed and their evaluation.  Model Development  1.Data Preparation:  - The dataset was cleaned to remove rows with missing values, focusing on relevant features.  - The target variable for prediction was the Draft Pick.  2. Splitting the Data:  - The dataset was divided into training and testing sets using an 80-20 split to ensure that the model could be trained effectively while retaining a portion of the data for evaluation.  3. Model Selection:  - A Linear Regression model was chosen due to its simplicity and interpretability, allowing for easy understanding of how each feature contributes to the prediction of draft pick positions.  4. Model Training:  - The model was trained using the training dataset, fitting it to learn the relationships between the features and the target variable.  Model Evaluation  1. Predictions:  - After training, predictions were made on the test dataset to evaluate how well the model performed.  2. Evaluation Metrics:  - Three primary metrics were used to assess model performance:  - Mean Squared Error (MSE): This metric measures the average squared difference between predicted and actual draft pick positions.  - R-squared (R²) Score: This statistic indicates how well the independent variables explain the variance in the dependent variable (draft pick). An R² score close to 1 indicates a good fit.  -Mean Absolute Error(MAE) : A lower MAE indicates better accuracy in predicting the draft pick. | | | | | | | | |  |
|  | Results-  - The model achieved an R² score of approximately 0.92, meaning about 92% of the variance in draft pick positions. This indicates the players have better success rate as for the better performance from the players in regards to their combine measurements.  Key insights and Conclusion**:**   * Player Measurements Matte**r**: Players with better vertical jump performance and faster sprint times tend to be drafted higher. * Height and Wingspan are strongly correlated, but height alone does not strongly predict draft success. * Height Advantage: Taller players generally have better chances of being drafted higher, which aligns with traditional scouting practices that favor size. * Weight Considerations: While weight can be an asset, excessively heavy players may face challenges during the selection process. Body Fat has a weak but notable relationship with draft pick, where players with lower body fat percentages tend to be drafted earlier.   The analysis underscores the importance of physical attributes such as height and vertical jump in determining a player's draft position. Teams should prioritize these metrics during evaluations while also considering the balance between size and athleticism for optimal player selection strategies. This comprehensive approach can enhance scouting effectiveness and improve team performance in future drafts.  Recommendations:  1. Focus on Athletic Metrics: Scouting teams should prioritize vertical jump and agility metrics when evaluating potential draftees as these factors significantly correlate with success.  2. Consider Height in Evaluations: Height should remain a key consideration in player evaluations since it correlates positively with draft success.  3. Balanced Player Profiles: Teams should look for a balance between size and athleticism when drafting players to ensure they meet both physical requirements and performance expectations. | | | | | | | | |  |
|  |  | | | |  |  | | | |  |
|  | |  | | | | |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  | Visualization of above analysis: |  | |
|  |  | |  |
|  |  |  | |







|  |  |  |
| --- | --- | --- |
|  |  |  |

