**Phase-1**

1. **Define Problem statement**

The Cricket ODI dataset, capturing information about One Day International (ODI) cricket matches, presents a valuable opportunity to perform comprehensive exploratory data analysis (EDA) and statistical analysis. The dataset includes details about ODI matches, such as the participating teams, match winner, venues, innings, and runs scored. The objective of this project is to conduct a thorough EDA and statistical analysis of the Cricket ODI dataset to gain insights into the factors affecting match outcomes and to draw meaningful conclusions. By examining the dataset, we aim to answer questions related to team performance, the impact of factors like venue and innings, and any patterns or trends that can be identified within the data. Project aims this analysis to provide valuable insights into the world of ODI cricket.

1. **Create project plan and product backlog**

**Project Plan:**

**Project Objective:** To perform an exploratory data analysis (EDA), conduct statistical analysis, and analyse probability distributions on a cricket One Day International (ODI) dataset to gain insights into player performance and match outcomes.

**1. Data Collection and Preprocessing :**

* Collect ODI dataset from reliable sources (e.g., ICC, ESPN, or Kaggle).
* Clean and preprocess data to handle missing values, data types, and outliers.

**2. Exploratory Data Analysis (EDA) :**

* Generate summary statistics .
* Create visualizations (bar charts, histograms, scatter plots) to represent key insights.
* Identify trends and correlations within the dataset.

**3. Statistical Analysis :**

* Conduct hypothesis testing to answer specific questions.
* Analyse batting averages, bowling averages, centuries, and other relevant cricket statistics.
* Perform t-tests, ANOVA, or other statistical tests to support hypotheses.

**4. Probability Distribution Analysis :**

* Analyse probability distributions for cricket metrics, including runs scored, and match results.
* Calculate probabilities, expected values, and confidence intervals.
* Visualize probability distributions (histograms, density plots).

**5. EDA and Statistical Analysis :**

* EDA and statistical analysis on the dataset.
* Include visuals, tables, and interpretations.
* Discuss key insights and trends.

**6. Probability Distribution Analysis Report**

* Probability distribution analysis .
* Present findings and insights from the analysis.

**Product Backlog:**

1.To implement the following statistical tests on the Cricket ODI dataset:

* Dispersion for the parameters
* Data distribution
* Test statistic
* Test type (T-test, Z-test, F-test, ANOVA, Chi-Square, PCA)

2.To implement the following data visualization techniques on the Cricket ODI dataset:

* Distribution
* Histogram
* Scatter plots
* Bar plot
* Box plot
* FacetGrid
* Heatmap
* Crosstab
* Line plot

3.To conduct the following exploratory data analysis on the Cricket ODI dataset:

* Univariate analysis
* Bivariate analysis
* Multivariate analysis

1. **Creation of Git Repository**

[**https://github.com/Shilpacs25/119CS21048-Mini-Project-**](https://github.com/Shilpacs25/119CS21048-Mini-Project-)