**README**

**EXPLANATION OF THE COLUMNS OF THE CLEANED AND TRANSFORMED DATA**

**'Date': Date of the race**

**'City': City of the race**

**'Race\_n\_time': Time and number of the race on the given date**

**'ID': ID**

**'Racer',: Racer of current race**

**'Racer\_1': Racer of ID latest race**

**'LAST\_RACER\_SAME\_AS\_CURRENT\_RACER':** TRUE IF **RACER** OF THE LAST RACE WAS SAME AS CURRENT RACE FOR THE GIVEN ID, False if not

**'LAST\_ RACER \_SAME\_AS\_CURRENT\_ RACER \_BINARY':** 1 if above column was TRUE, False if not

**'Yas':** Age of the given ID

**'Speed\_in\_Secs':** Speed of the ID in the given race converted in m/s

**'GANYAN\_TRANS':** Odds of the ID for the given race

**'GANYAN\_TRANS\_int':** Odds in the above column converted

**'RANK':** Rank of the ID in the given race

**'AG\_Percentage':** Percentage (Only available for specific races)

**'AG\_RANK\_AMONG\_GROUP':** Rank of the percentage number of AG for the given race

**'AG\_RANKING\_Score\_Normalized':** Min-max normalized score of AG calculated from the rank of the AG

**'Ranking\_Score\_Normalized':** Min-max normalized score from the rank of the ID for the given race

**'First\_Position\_Percentage':==1A=**Percentage of the 1st Position of the Racer of the ID for current race

**'Second\_Position\_Percentage':==1B=**Percentage of the 2nd Position of the Racer of the ID for current race

**'Third\_Position\_Percentage':==1C=**Percentage of the 3rd Position of the Racer of the ID for current race

**'Forth\_Position\_Percentage':** Percentage of the 4th Position of the Racer of the ID for current race

**'TOP\_THREE\_RACER\_PERCENTAGE': 1A+1B+1C**

**'RANK\_LATEST\_1ST':** RANK OF THE ID IN THE LATEST RACE

**'RANK\_LATEST\_2ND':** RANK OF THE ID IN THE 2ND LAST RACE

**'RANK\_LATEST\_3RD':** RANK OF THE ID IN THE 3RD LAST RACE

**'RANK\_LATEST\_4TH':** RANK OF THE ID IN THE 4TH LAST RACE

**'RANK\_LATEST\_5TH':** RANK OF THE ID IN THE 5TH LAST RACE

**'TOPLAM\_RACES':** Total Races played by the ID in whole lifetime

**'TOPLAM\_1ST':** Number of races in which ID got 1st position

**'TOPLAM\_2ND':** Number of races in which ID got 2nd position

**'TOPLAM\_3RD':** Number of races in which ID got 3rd position

**'TOPLAM\_4TH':** Number of races in which ID got 4th position

**'BIRINCILIK\_PERCENTAGE':** Percentage of 1st Positions

**'IKINCILIK\_PERCENTAGE':** Percentage of 2nd Positions

**'UCUNCULUK\_PERCENTAGE':** Percentage of 3rd Positions

**'DORDUNCULUK\_PERCENTAGE':** Percentage of 4th Positions

**'SPEED\_LATEST\_1ST':** Speed of the last race of the given ID calculated from the time and distance for that specific race

**'SPEED\_LATEST\_2nd':** Speed of the 2nd last race of the given ID calculated from the time and distance for that specific race

**'SPEED\_LATEST\_3rd':** Speed of the 3rd last race of the given ID calculated from the time and distance for that specific race

**'SPEED\_LATEST\_4th':**Speed of the 4th last race of the given ID calculated from the time and distance for that specific race

**'SPEED\_LATEST\_5th':** Speed of the 5th last race of the given ID calculated from the time and distance for that specific race

**'RANK':** Rank of the ID in the given race, **WILL BE OUR LEARNING CRITERIA**

**'Ranking\_Score\_Normalized':** Min-max normalized score from the rank of the ID for the given race, **WILL BE OUR PREDICTION CRITERIA OR SCORE**