

# TOUR DE FRANCE DASHBOARD



Country of Winner  
All

Average Winner Height

1.78 m

Average Winner Weight

68.37 kg

Average Winner Speed

34.55 km/h

Number of Races

102

Average Finish Time

07:19:01

Number of Stages

21

Number of Starters

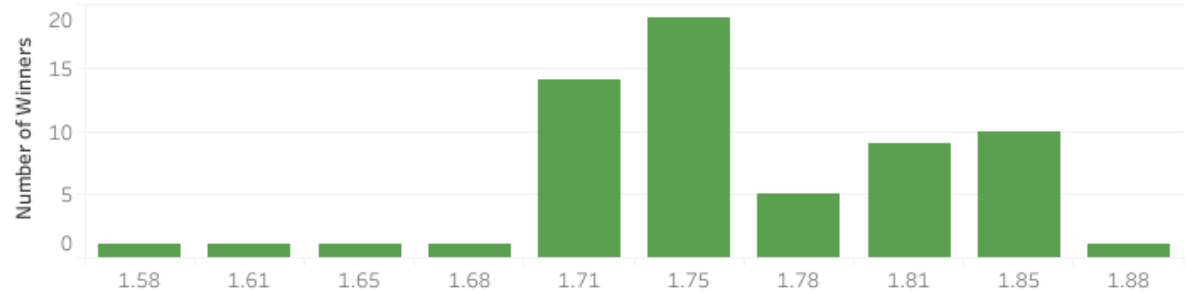
15,707

Number of Finishers

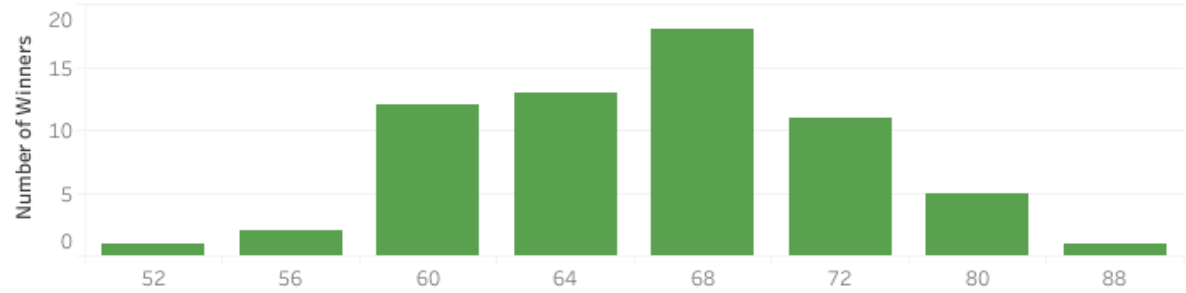
9,895

Team of Winner  
All

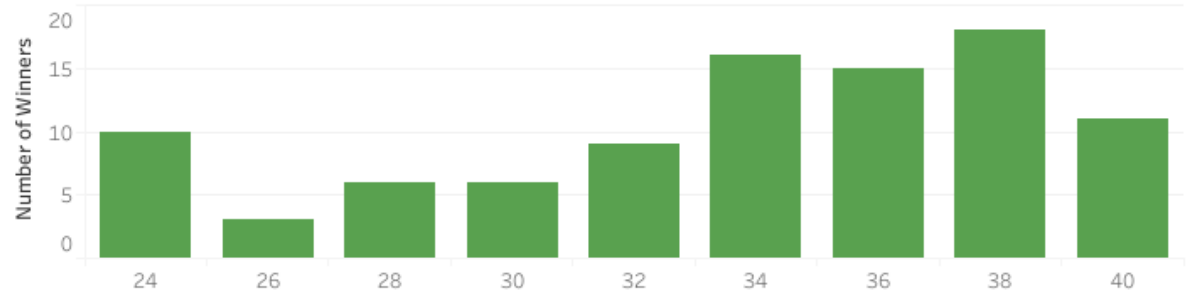
Winner Heights (m)



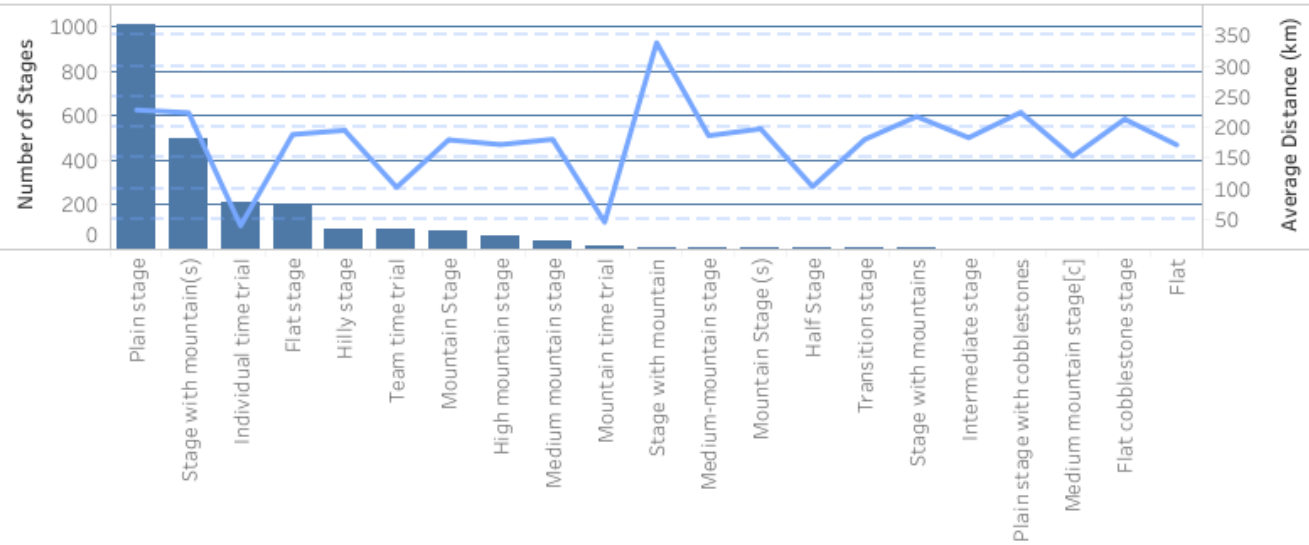
Winner Weights (kg)



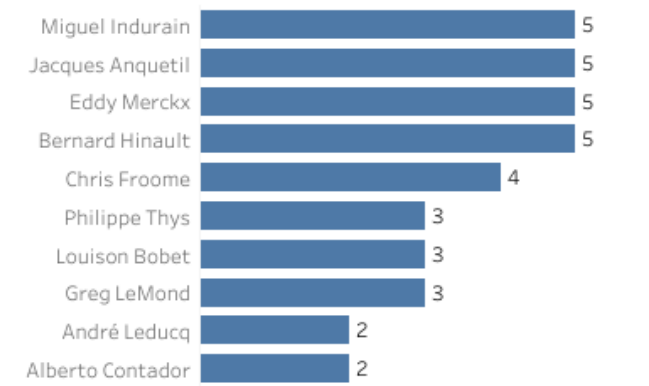
Average Winner Speeds (km/h)



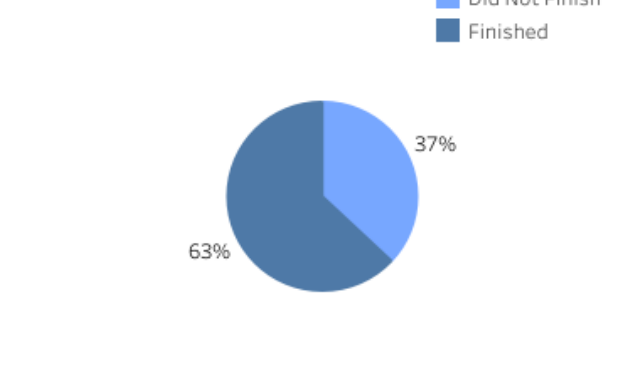
Stage Breakdown



Top 10 Winners



Race Results



# Abstract

- ▶ I assumed the role of a data visualization expert at the Amaury Sport Organization
  - ▶ This company organizes the annual Tour de France cycling race.
- ▶ My goal is to increase the race's popularity around the world by creating an infographic-style visual that will do the following:
  - ▶ Educate new viewers
  - ▶ Highlight the magnitude of the event
  - ▶ Build anticipation for this year's race
- ▶ To accomplish this goal, I created an exploratory dashboard that summarizes key insights into a historic dataset for every Tour de France from 1903-2022.
  - ▶ The dashboard is interactive.
  - ▶ The user can filter out the country and teams of each Tour de France winner.
- ▶ I have uploaded all the files for this project onto my [GitHub](#).

# Data Cleaning Strategy

- ▶ The dataset was a combination of structured and unstructured data.
  - ▶ Therefore, my data cleaning strategy was a little bit involved.
- ▶ This was how I cleaned the “tdf\_finishers” table:
  - ▶ Split the Country and the Rider into two separate columns.
  - ▶ Changed the Time values to duration arrays.
- ▶ This was how I cleaned the “tdf\_stages” table:
  - ▶ Removed the imperial values for the distances.
  - ▶ Split the Country and Winner into two separate columns.
  - ▶ Split the Starting and Ending Cities of the Courses into two separate Columns.
- ▶ This was how I cleaned the “tdf\_tours” table:
  - ▶ Removed the imperial values for the distances.
  - ▶ Combined Starters and Finishers into one column.
- ▶ This was how I cleaned the “tdf\_winners” table:
  - ▶ Converted Time values to duration arrays.
  - ▶ Removed the units from the Height, Weight, and Avg Speed and converted the values to numeric data types.
    - ▶ I placed the units for each metric in their corresponding column name.

# Data Model



# Insights

- ▶ Here are the following insights into the Tour de France Winners:
  - ▶ Most frequent winners:
    - ▶ Miguel Indurain (5 wins)
    - ▶ Jacques Anquetil (5 wins)
    - ▶ Eddy Merckx (5 wins)
    - ▶ Bernard Hinault (5 wins)
  - ▶ Most frequent winner height range: 1.75 - 1.78 m
  - ▶ Most frequent winner weight range: 68 - 72 kg
  - ▶ Most frequent winner average speed range: 38 - 40 km/h
- ▶ Here are the following insights into the Tour de France:
  - ▶ Race completion rate: 63%
  - ▶ Most frequent stage: Plain Stage (1,008 stages)
  - ▶ Stage with the longest average distance: Stage With Mountain (337.8 km)