TOUR DE FRANCE DASHBOARD

Average Winner Height	Average Winner Weight	Average Winner Speed	Number of Races	Average Finish Time	Number of Stages	Number of Starters	Number of Finishers
1.78 m	68.37 kg	34.55 km/h	102	07:19:01	21	15,707	9,895

Winner Heights (m)
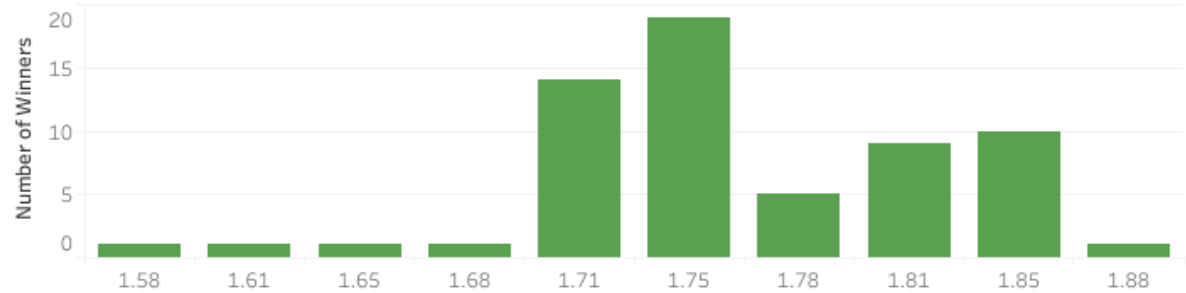
Stage Breakdown

Winner Weights (kg)

Average Winner Speeds (km/h)

Top 10 Winners

| Winner	
Miguel Indurain	5
Jacques Anquetil	5
Eddy Merckx	5
Bernard Hinault	5
Chris Froome	4
Philippe Thys	3
Louison Bobet	3
Greg LeMond	3
André Leducq	2
Alberto Contador	2

Race Results

Did Not Finish
Finished

37%

63%
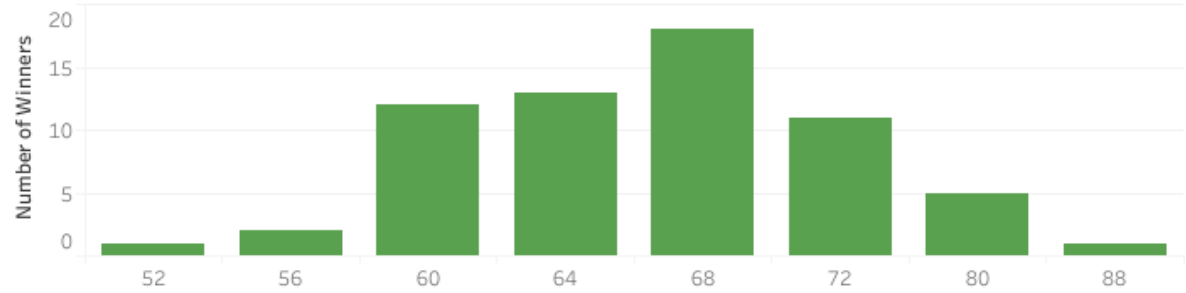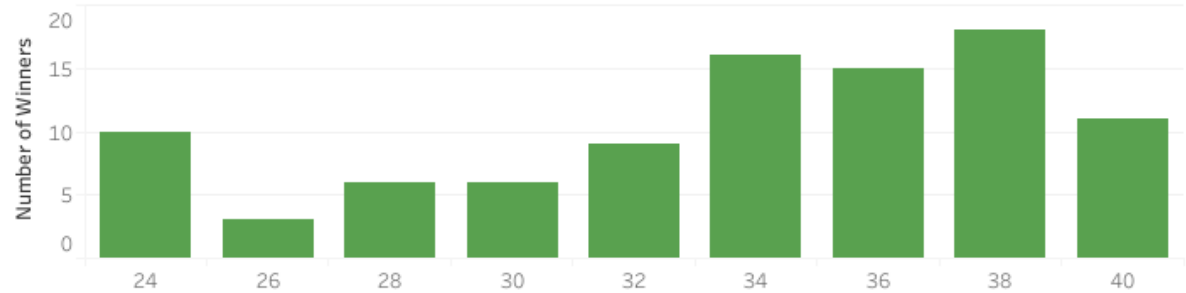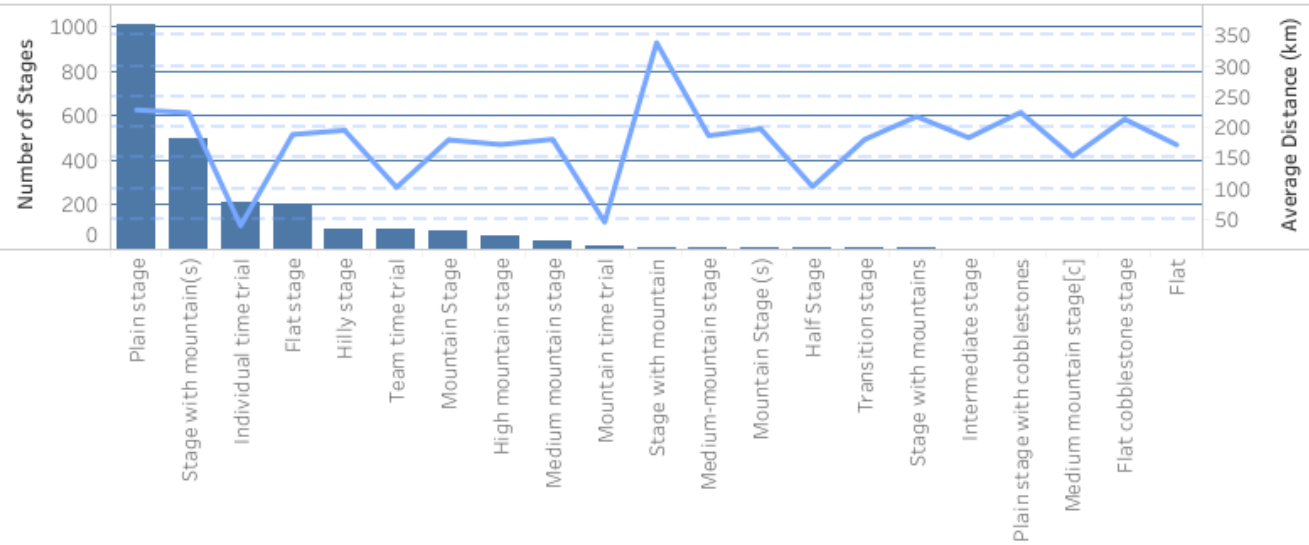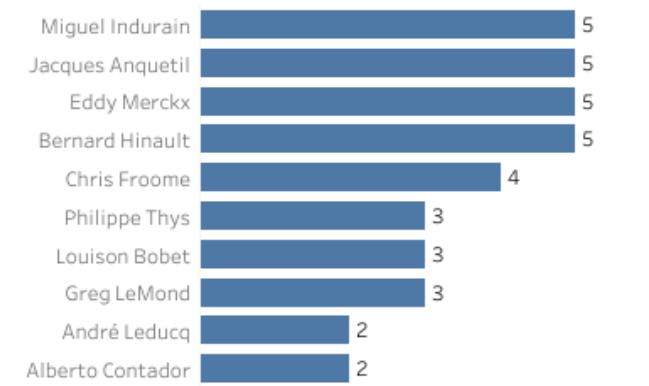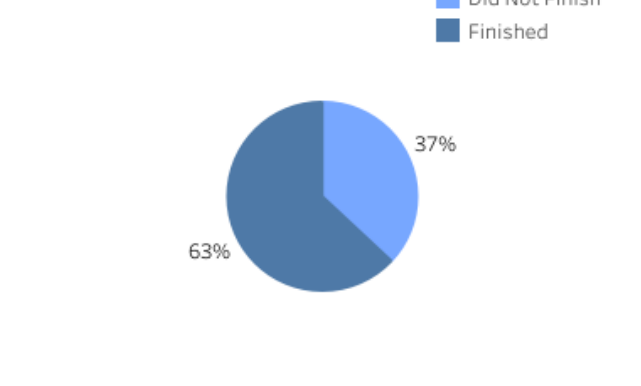
Abstract

- 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)