



Predicting Cycling Metrics

by Adriana J. Machado

Table of Contents



01

**Background &
Problem Statement**

02

**Data Gathering &
Extraction**

03

**Variables & Feature
Engineering**

04

**Exploratory Data
Analysis**

05

Modeling

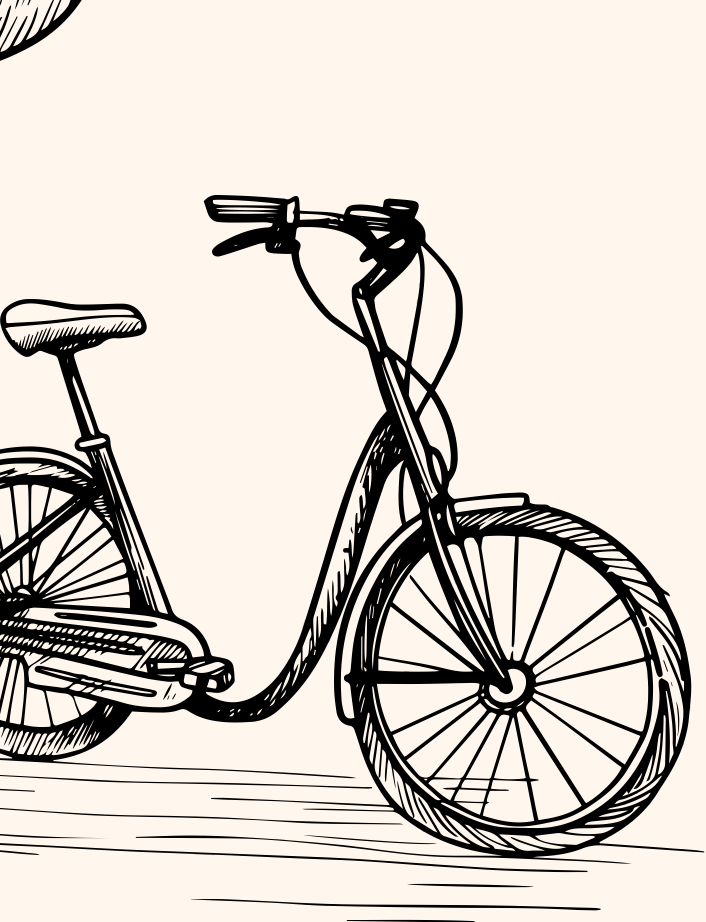
06

**Tableau Dashboards &
Streamlit**



01

Background & Problem Statement



Germany 1817

First verified account of the bicycle, “draisine”

\$6,200,000,000 USD

2004-2005 US bicycle market

52.73 billion

Estimated cyclists in the USA in 2020

Training for a Century & Not Training for a Century

**April '22
High**

Peak training mode

Cycling >100 miles per week

**July '22
Average**

Mostly cycling for regular exercise

<30 miles per week



Cycling is a widely popular sport and can accommodate a range of experiences. Being able to predict how much stress a route may put on your body can be really useful and offer a lot of insight for goal-making or reality-checking.

Given past periods of cycling performance, how will I perform on a given route, based on heart rate intensity zones? Is that trail going to be too difficult or am I capable of riding it without feeling like I'm going to die?



02

Data Gathering & Extraction



Data Sources Used



Garmin GPX Files

GPS Exchange

Trackpoint data w/
biometrics



OpenWeather API

Historical
Current
Forecast

03

Variables & Feature Engineering



Predicting Heart Rate



Location & Distance

Latitude
Longitude
Distance
between points

Total time
Trackpoint
differences

Time



Elevation

Elevation
changes

Temperature,
humidity, wind,
etc.

Weather



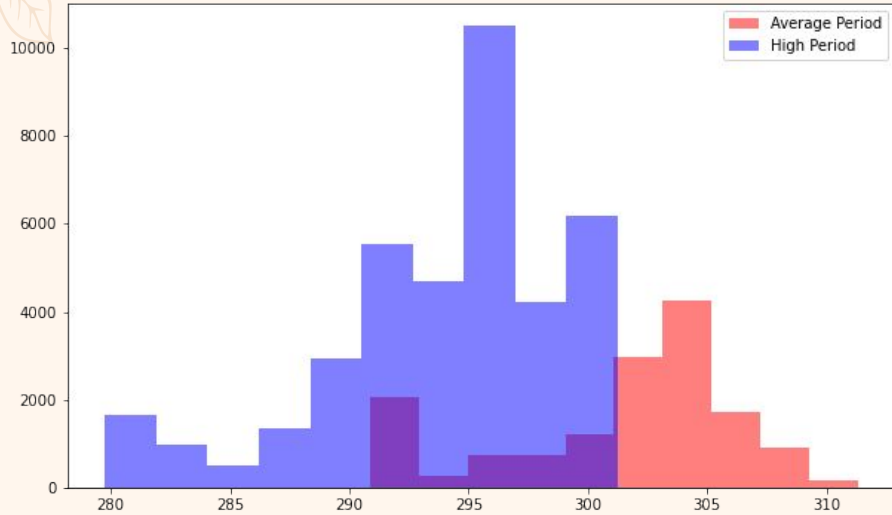
04

Exploratory Data Analysis

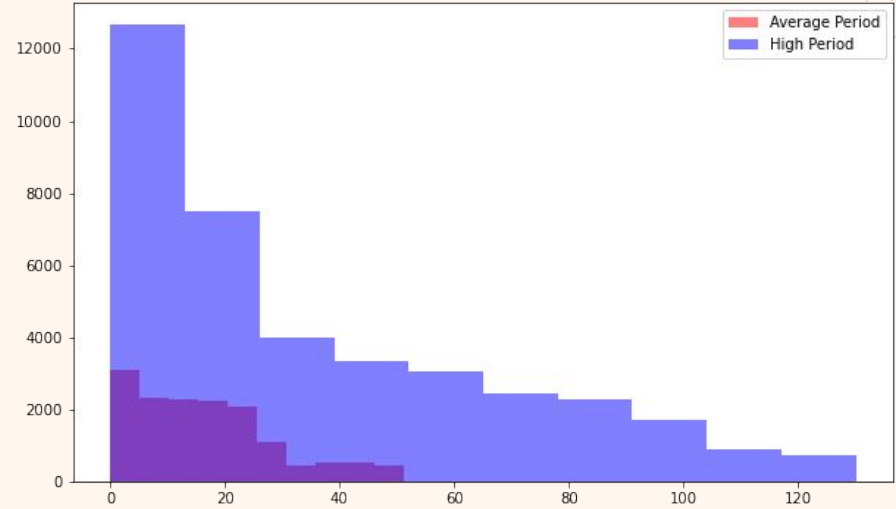


Major Differences

Temperature(K) Histograms

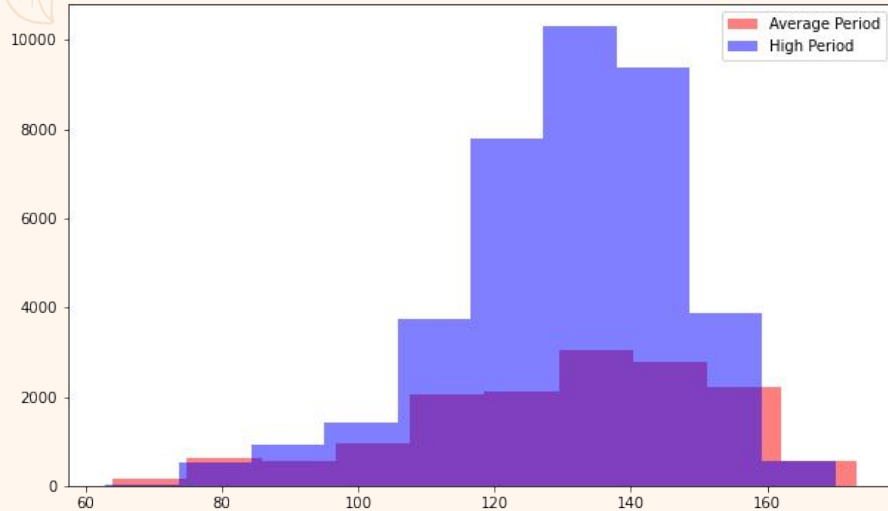


Distance(km) Histograms

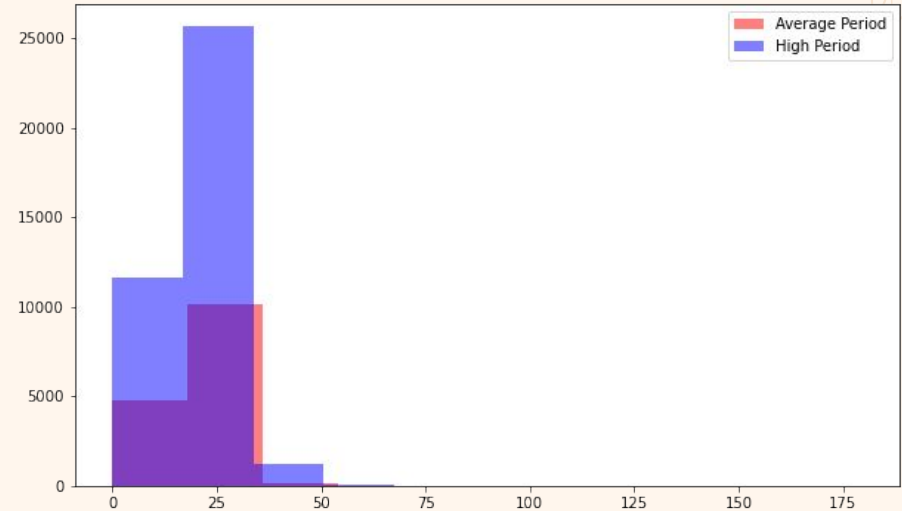


Surprising Similarities

Heart Rate Histograms

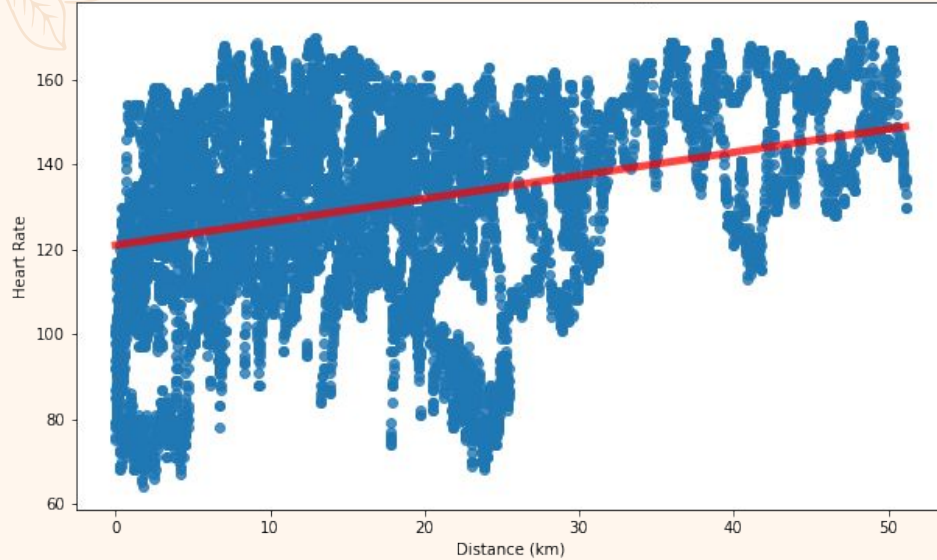


Speed(km/hr) Histograms

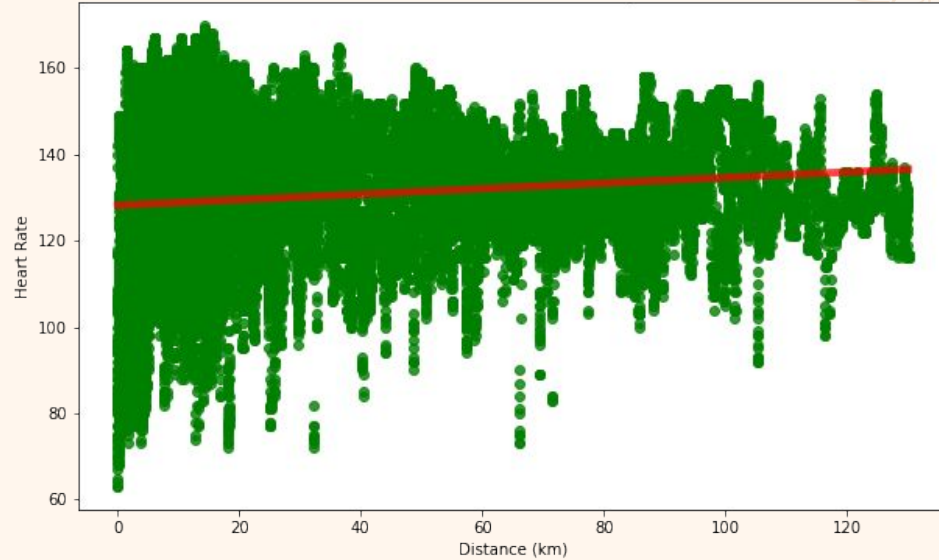


Average v High

Distance v Heart Rate - Average

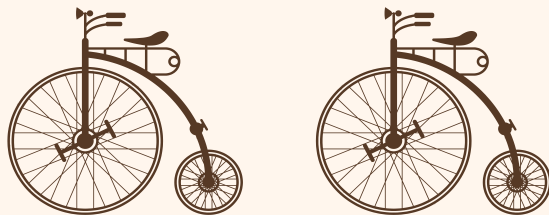


Distance v Heart Rate - High



05

Modeling



Regression Models

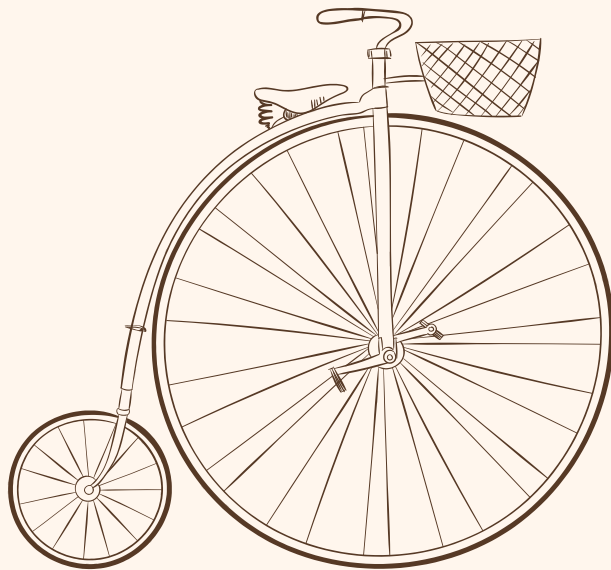
Neural Nets

**AdaBoosted
RandomForest**

XGBoost

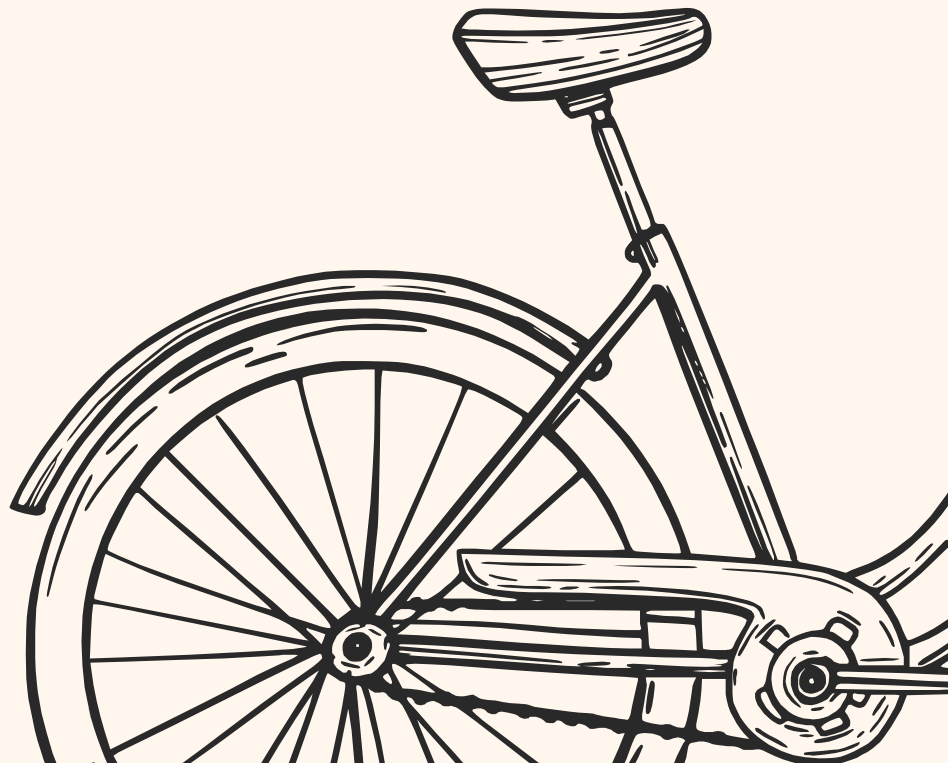


STREAMLIT



06

Conclusions & Recommendations



**I wouldn't have
died**



More variables?

Power output



**Personalized results
per user**

Thanks

Do you have any questions?

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

