

# ABSENTEEISM AT WORK

---

## Assignment 2 - Data Modeling & Presentation

---

Reuben Rajeev - s3717497  
Jeremy Quintana - s3719476

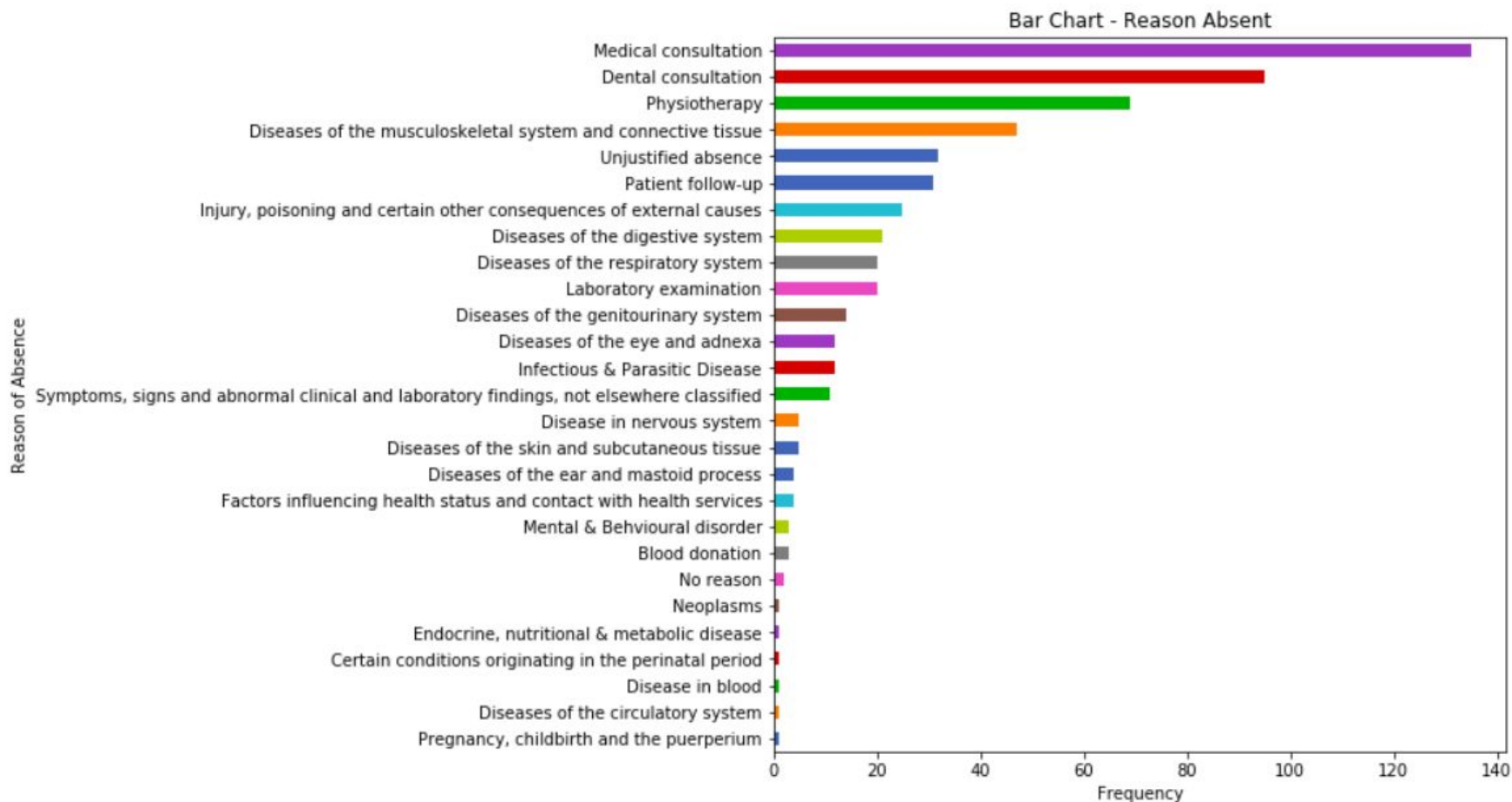
# DATASET & GOALS

## Dataset:

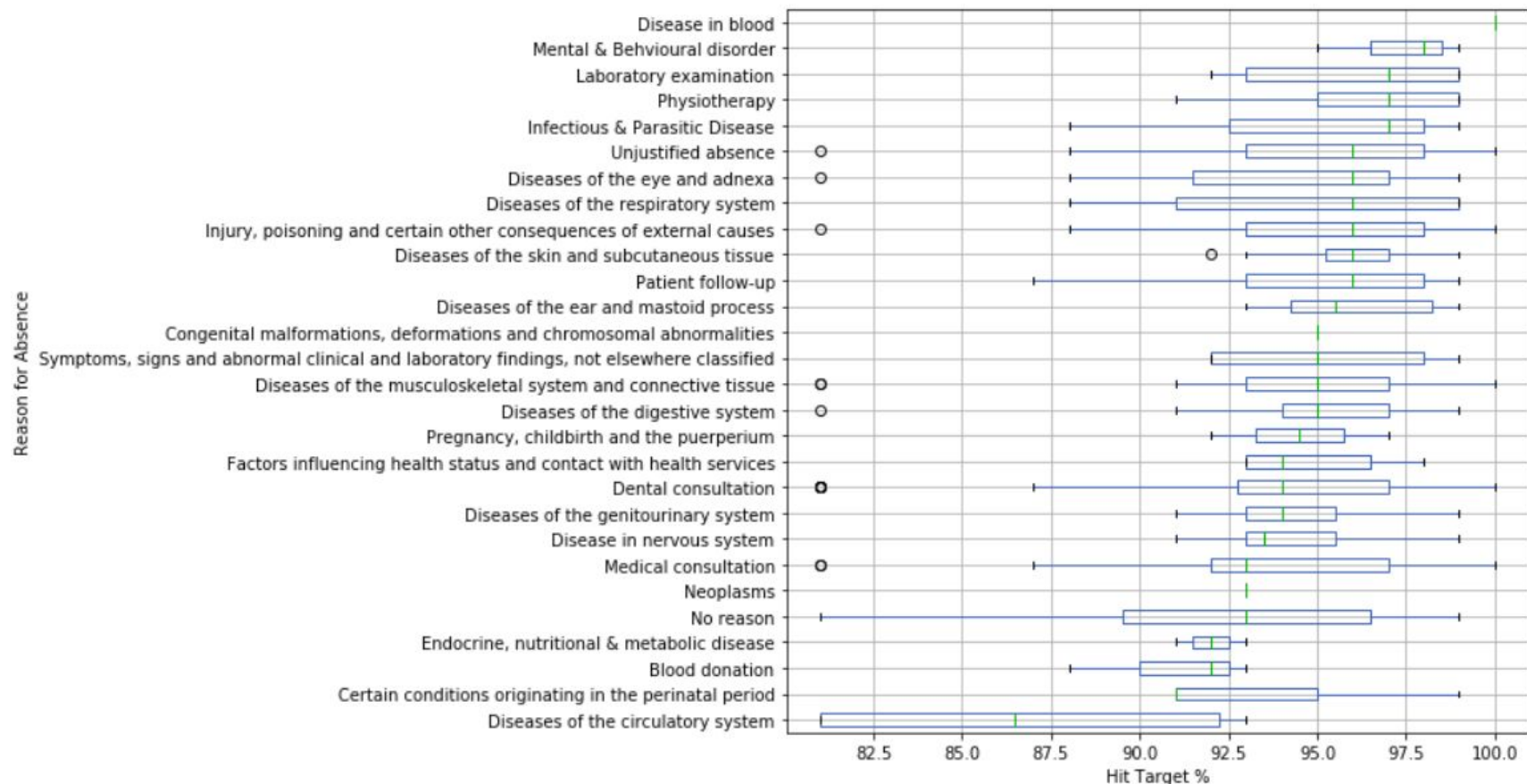
- Worker absences in a courier company
- 800 observations
- Absentee observations:
  - Characteristics
  - Reason for absence

## Goals:

- Company affects absences?
- ↓ absences ↑ profit



Boxplot of ordered reasons for absence for the hit target



# MODEL ERROR RATE

	Pre-Feature Engineering	Post-Feature Engineering
Classification Error Rate	63.54%	60.42%

	0.5	0.4	0.2
KNN	60.42%	58.87%	52.59%
Decision Tree	56.25%	52.81%	50.00%

# CONCLUSION & RECOMMENDATION