

**Data understanding:** our data comes from SDOT Traffic Management Division, Traffic Records Group in CSV format (**Data-Collisions.csv**). We have data about track conditions, position, severity, peoples involved, etc. All the columns and it's meaning is described in the **Metadata.pdf** file. We have an imbalanced dataset and many missing values, so we need to performe a data cleaning to have better accuracy in our models.