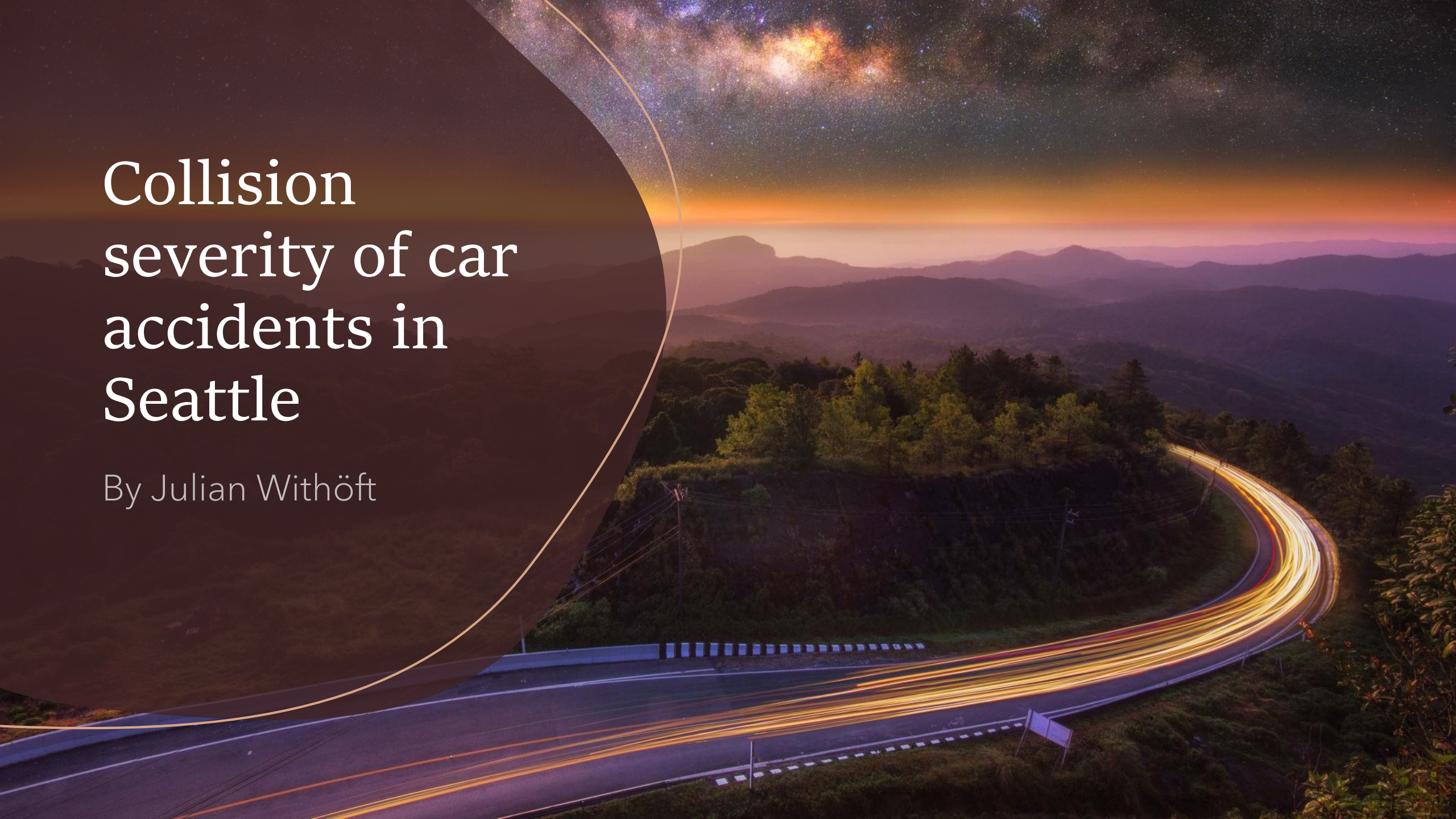
Collision severity of car accidents in Seattle

By Julian Withöft



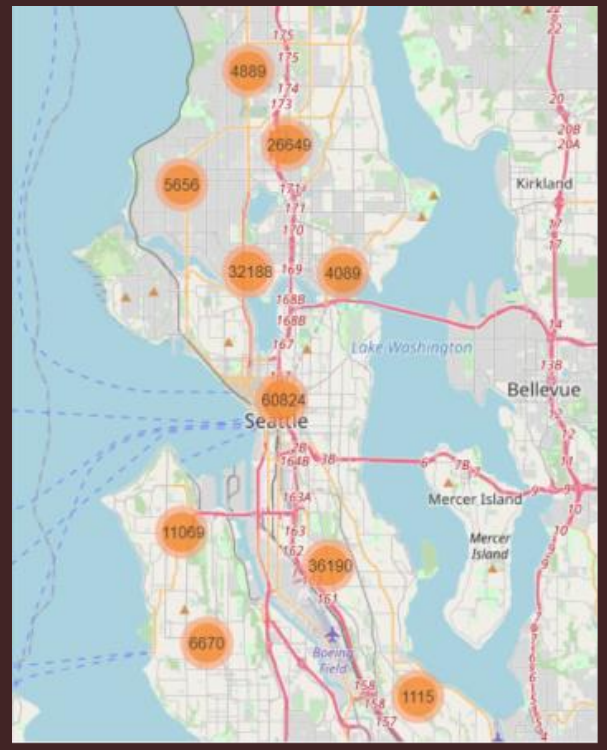
Motivation

- Making roads safer for everyone
- Predict & prevent accidents
- Distribute first aid responders intelligently

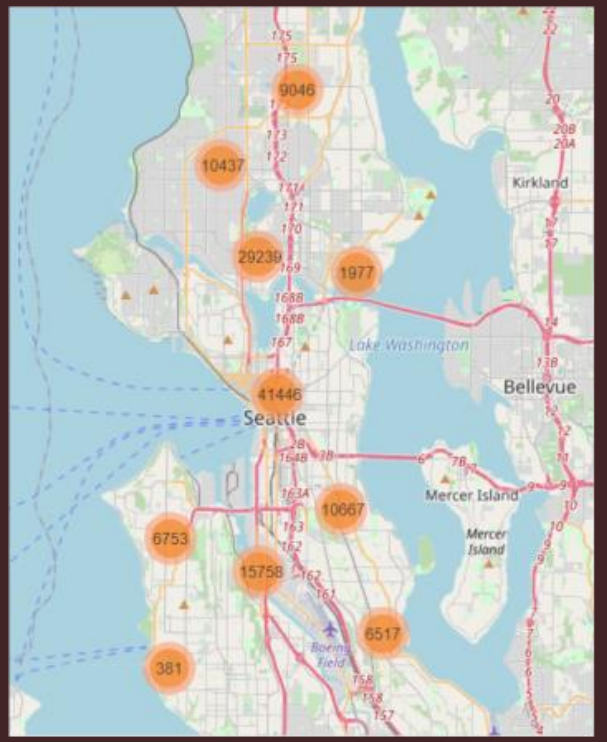
Data

- Real dataset of 194673 accidents that occured in Seattle
  - Provided in the IBM Data Science Capstone course
- Deletion of NaN values as well as Relabeling and Aggregating of similar entries
- Dependent variable: collision severity code
  - 1 – Vehicle collision; 2 – Human collision

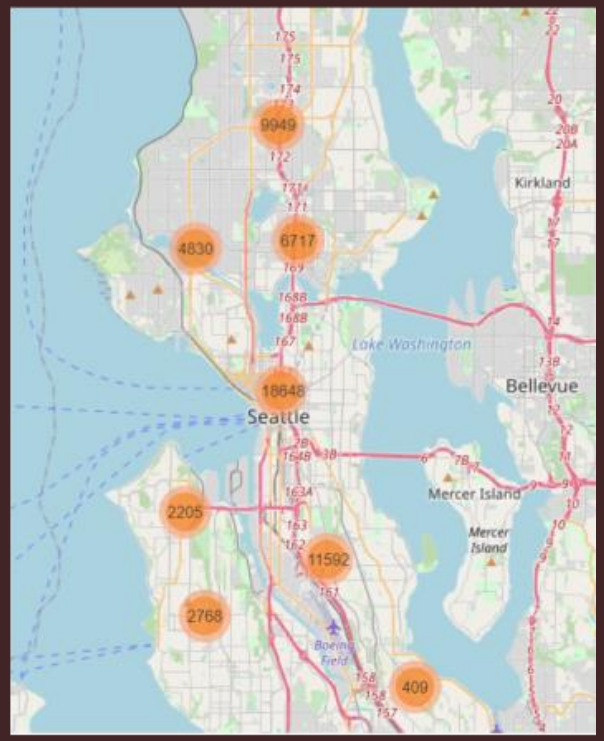
Geographical information



All accidents



Collision severity code 1



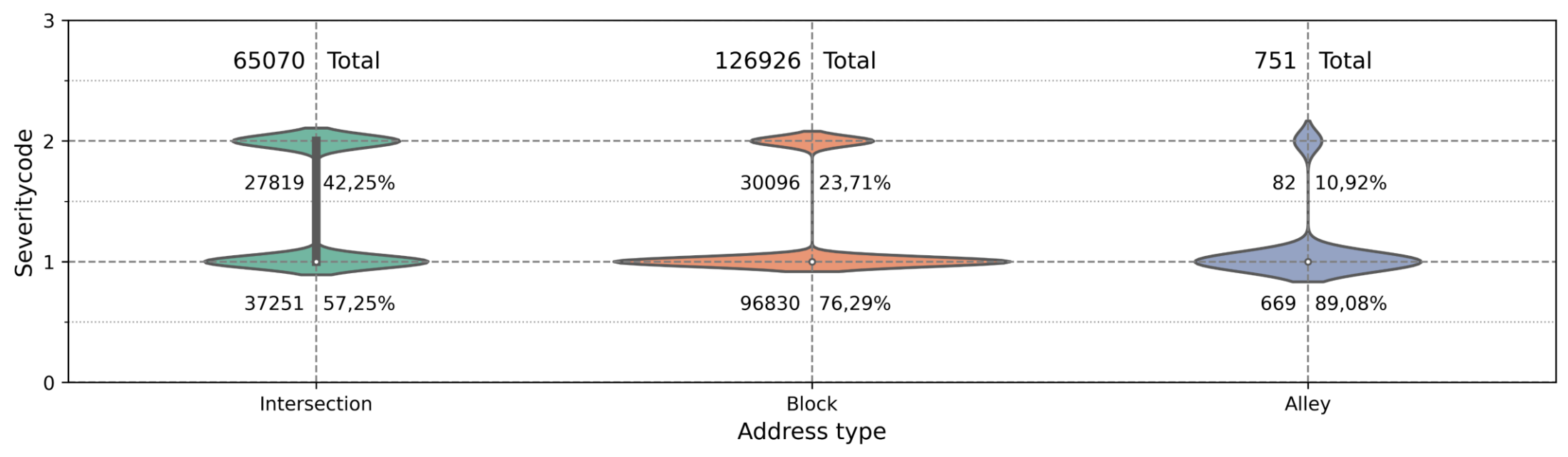
Collision severity code 2

- More accidents in the city center than in the suburban areas

➢ Dense and stimulated traffic in the city center

Address information

- Most human collisions at intersections
- ➢ Simultaneous crossing of pedestrians and turning of vehicles
- Alley with least human collisions, while block in between



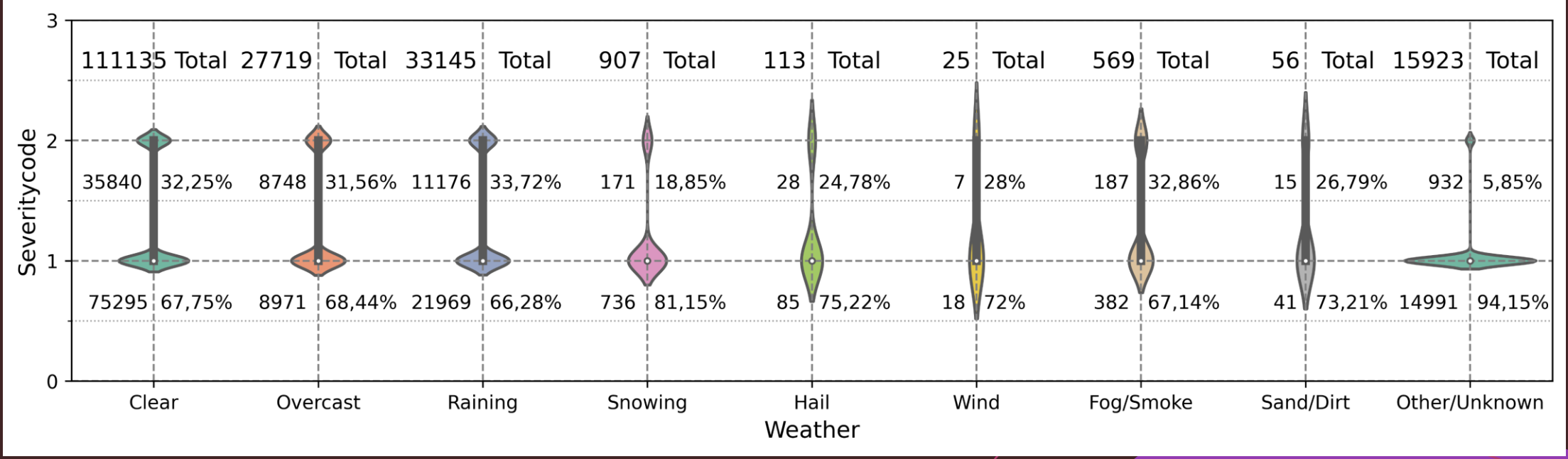
Weather information

- Snowing & Hail with less human collisions

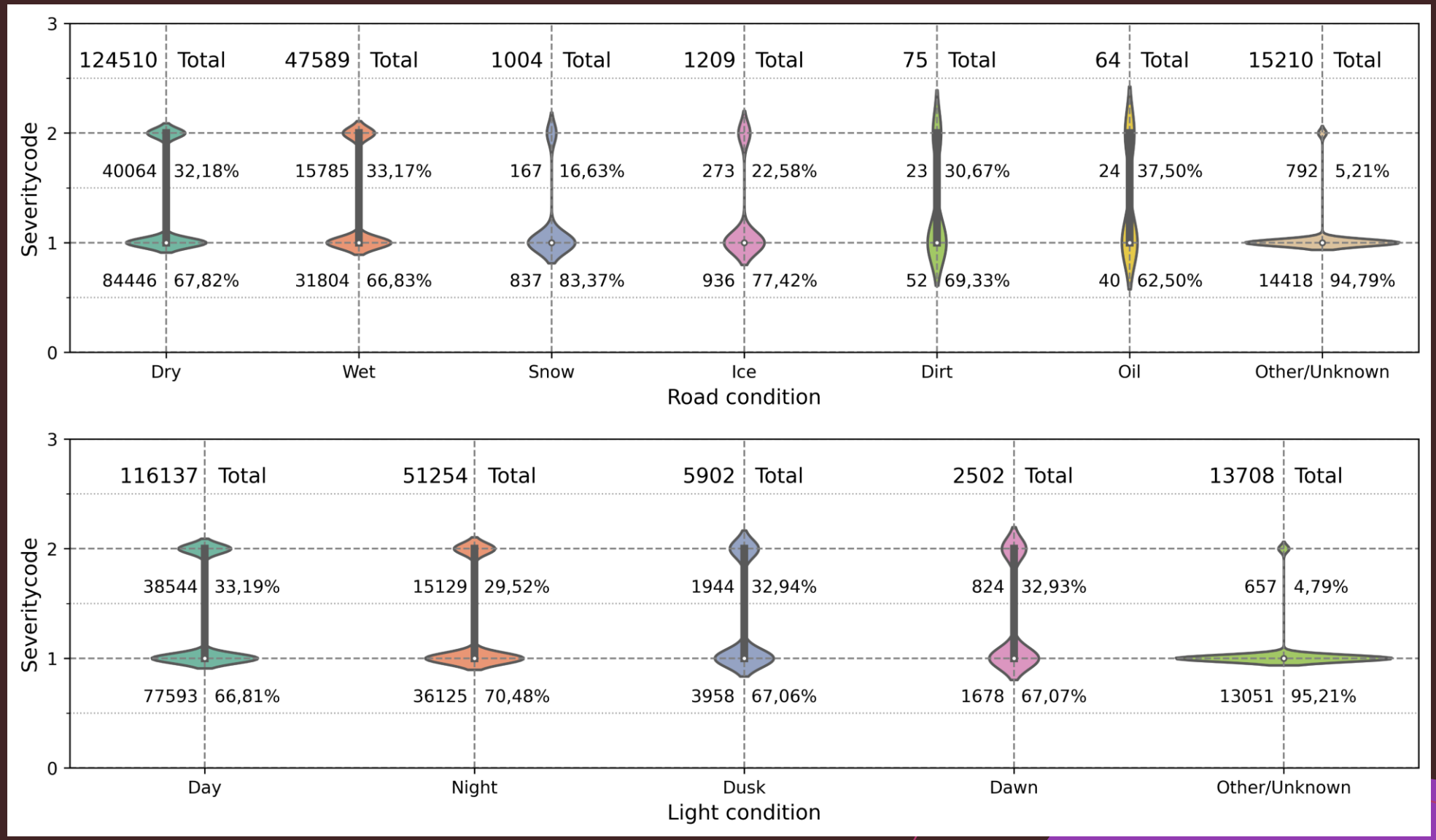
➢ Less pedestrians on the streets

- Fog/Smoke with more human collisions

➢ Strongly impaired visibility conditions

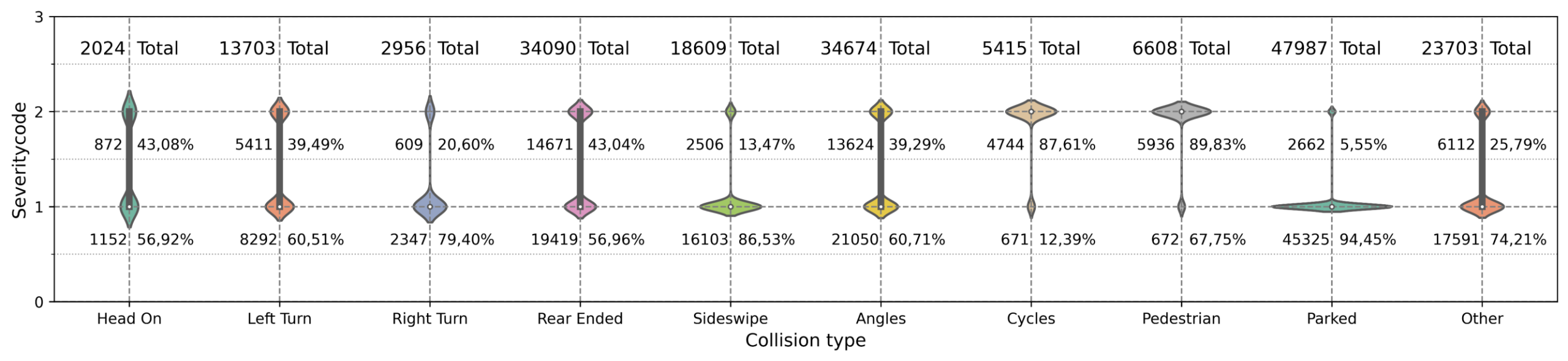


Road & light condition



Collision type & accident course information

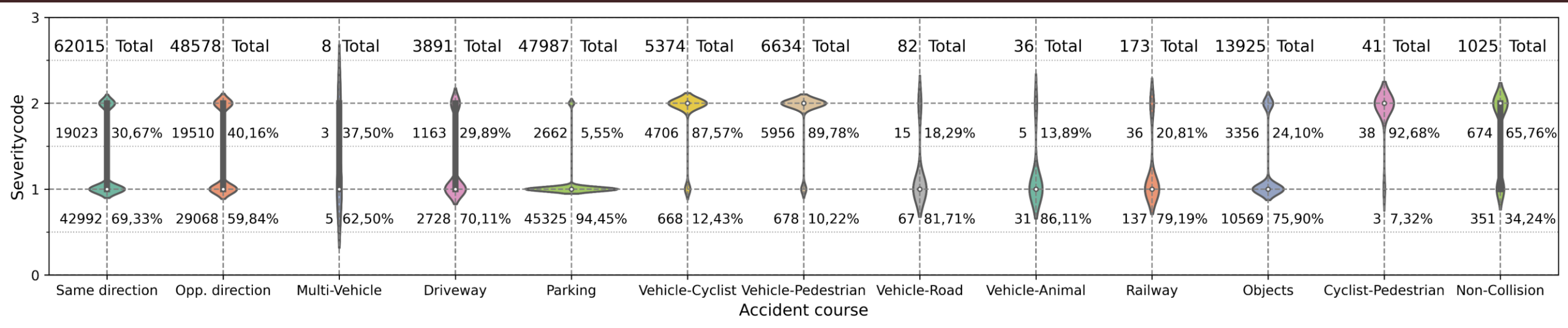
- Cycles & Pedestrians obviously enhance human collision
- Parked vehicles lead to very low human collision as they are usually unmanned



Accident course information

- Human collision for opposite direction accidents more likely than for same direction accidents

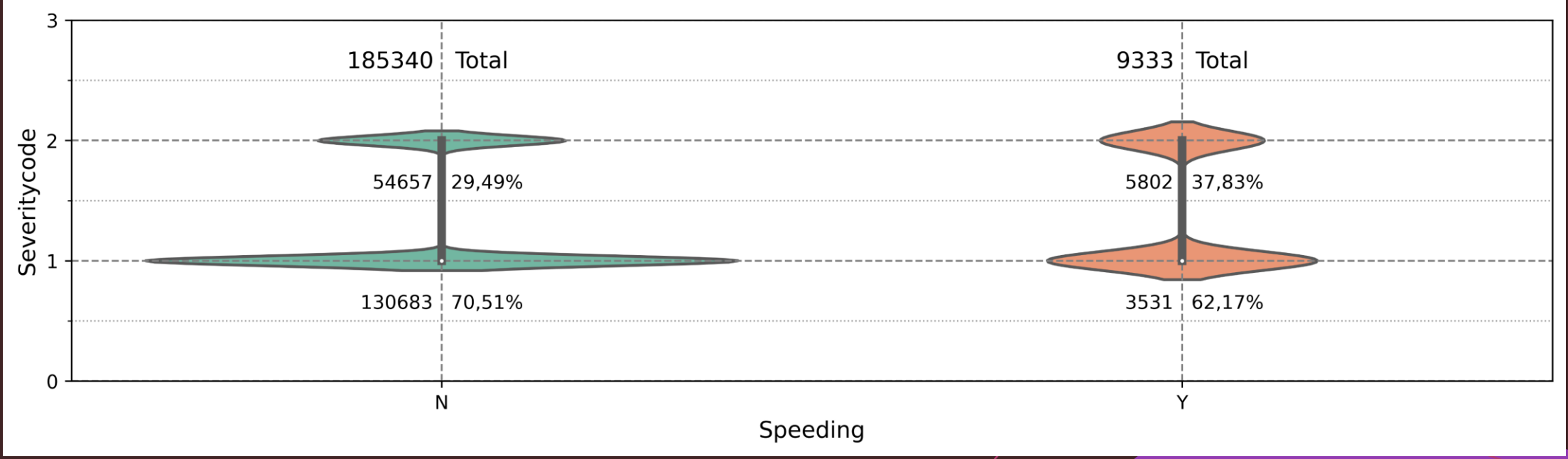
➢ Contrary forces generally increase the severity of the accident



Speeding information

- Increased velocity and less controllable maneuvers for the driver

➢ Human collision more likely to occur



Predictive modelling – Classification machine learning

- 4 classification algorithms: Decision tree, K-nearest neigbhor (kNN), Logistic regression and Support vector machine (SVM)

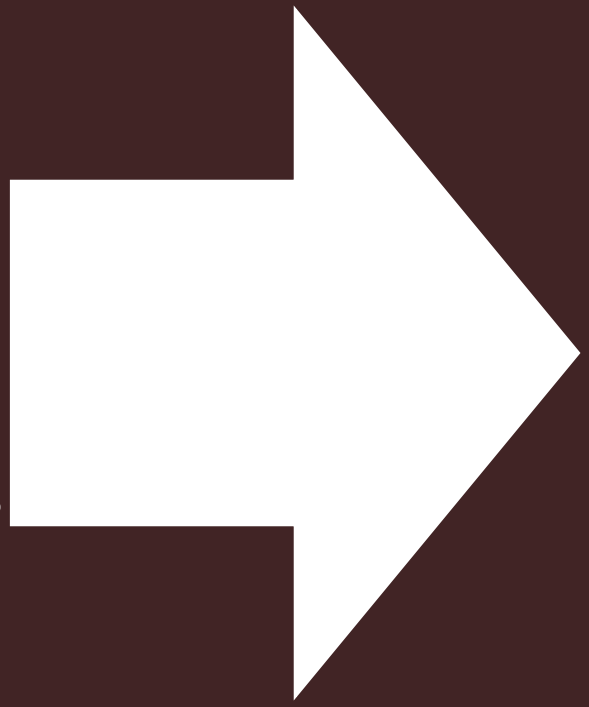
- Accuracy:

Decision tree: 75,25%

kNN: 75,31%

Log. regression: 74,97%

SVM: 75,81%

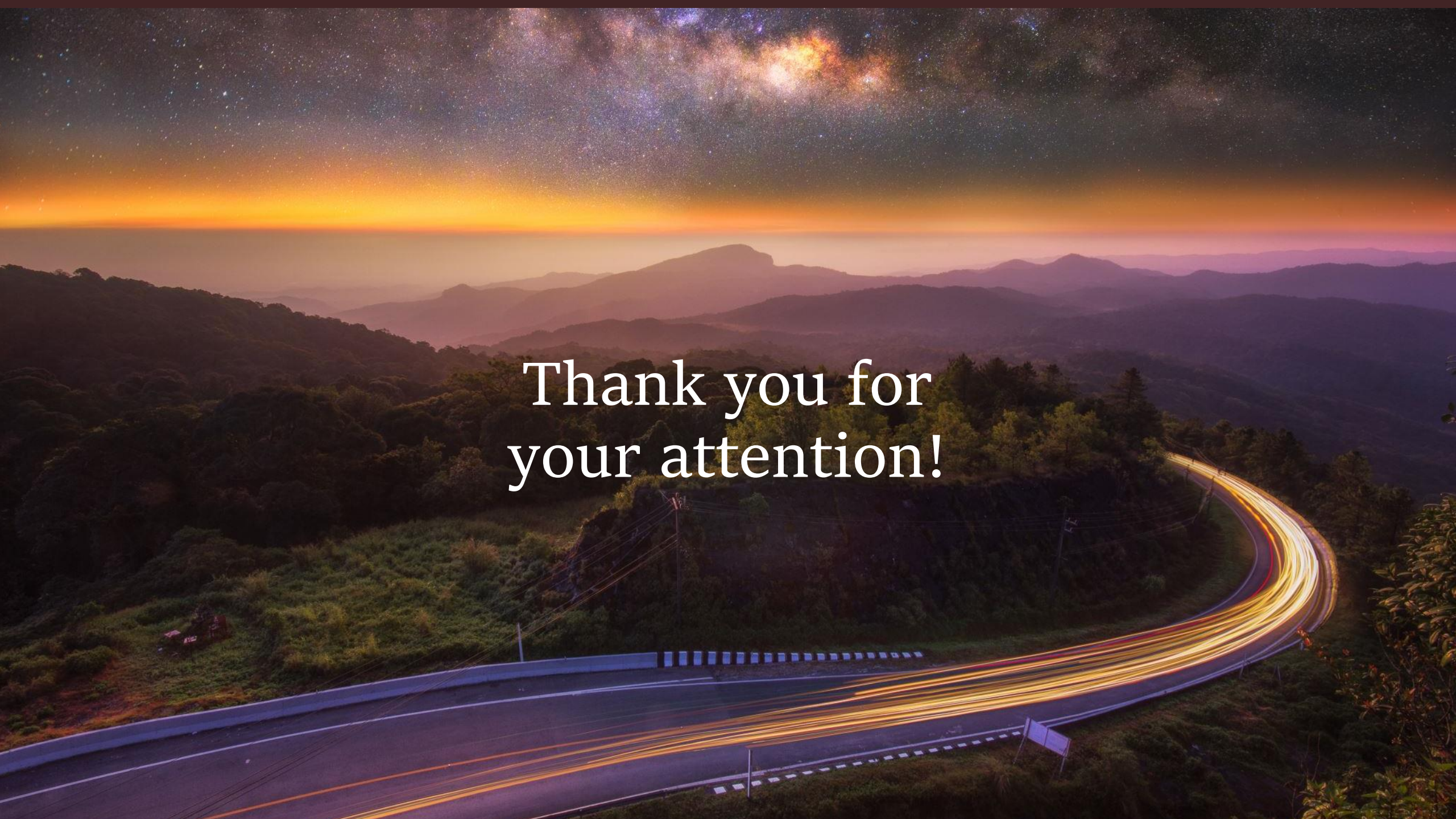


- SVM delivers best results
- Log. regression with the most inaccurate results

Conclusion

- Collision severity of car accidents in Seattle was analyzed
- Influence factors were analyzed statistically and graphically
- Development of machine learning classification algorithms

➢ Enables Pre-Classification of accidents (Predict & Prevent) and intelligent distribution of first aid responders to the accidents


Thank you for
your attention!