# Predicting the car accident rate

Background

▶ Car accidents happen frequently → disastrous impact

▶ The situation still cannot improve

  ▶ Which location has the highest severity of accidents?

  ▶ Under what weather conditions, will the accidents occur frequently?

  ▶ Under which combination of factors, will severity of accidents be the highest?

Data acquisition and cleaning

- **Data source:** Data-Collisions.csv

- **Data cleaning:** replace the missing value with NaN
  - Replace the categorial variable to numeric value

Exploratory Data Analysis

Weather condition and severity
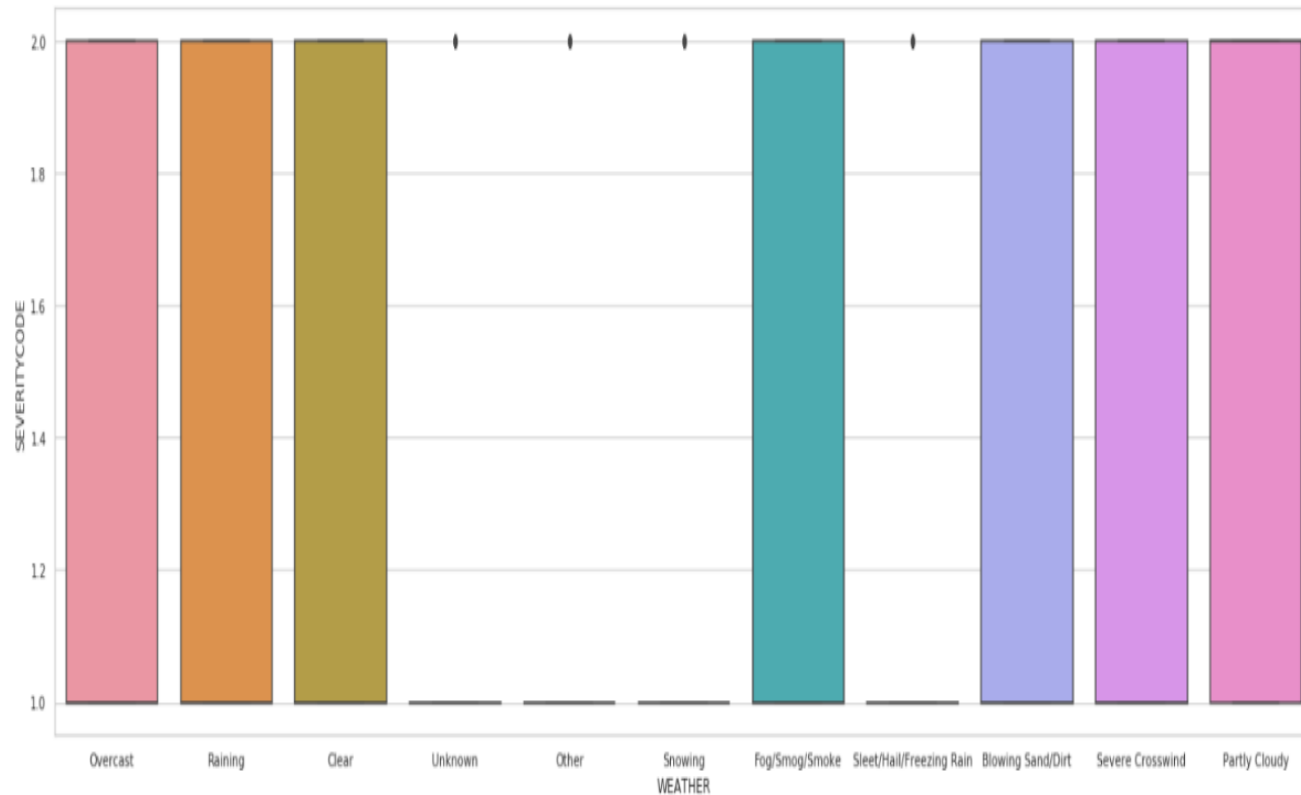
	WEATHER
Clear	111135
Raining	33145
Overcast	27714
Unknown	15091
Snowing	907
Other	832
Fog/Smog/Smoke	569
Sleet/Hail/Freezing Rain	113
Blowing Sand/Dirt	56
Severe Crosswind	25
Partly Cloudy	5

Road condition and accident

	ROADCOND
Dry	124510
Wet	47474
Unknown	15078
Ice	1209
Snow/Slush	1004
Other	132
Standing Water	115
Sand/Mud/Dirt	75
Oil	64

Light condition and accident

	LIGHTCOND
Daylight	116137
Dark - Street Lights On	48507
Unknown	13473
Dusk	5902
Dawn	2502
Dark - No Street Lights	1537
Dark - Street Lights Off	1199
Other	235
Dark - Unknown Lighting	11

Combination of factors

	WEATHER	ROADCOND	JUNCTIONTYPE	SPEEDING	ST_COLCODE	SEVERITYCODE
1	Blowing Sand/Dirt	Snow/Slush	Mid-Block (not related to intersection)	Y	50	2.0
11	Clear	Dry	At Intersection (but not related to intersection)	Y	13	2.0
18	Clear	Dry	At Intersection (intersection related)	Y	0	2.0
19	Clear	Dry	At Intersection (intersection related)	Y	1	2.0
28	Clear	Dry	At Intersection (intersection related)	Y	25	2.0
33	Clear	Dry	At Intersection (intersection related)	Y	30	2.0
35	Clear	Dry	At Intersection (intersection related)	Y	45	2.0
39	Clear	Dry	At Intersection (intersection related)	Y	73	2.0
41	Clear	Dry	At Intersection (intersection related)	Y	0	2.0
42	Clear	Dry	At Intersection (intersection related)	Y	1	2.0
49	Clear	Dry	At Intersection (intersection related)	Y	2	2.0
50	Clear	Dry	At Intersection (intersection related)	Y	21	2.0
53	Clear	Dry	At Intersection (intersection related)	Y	24	2.0

Modeling
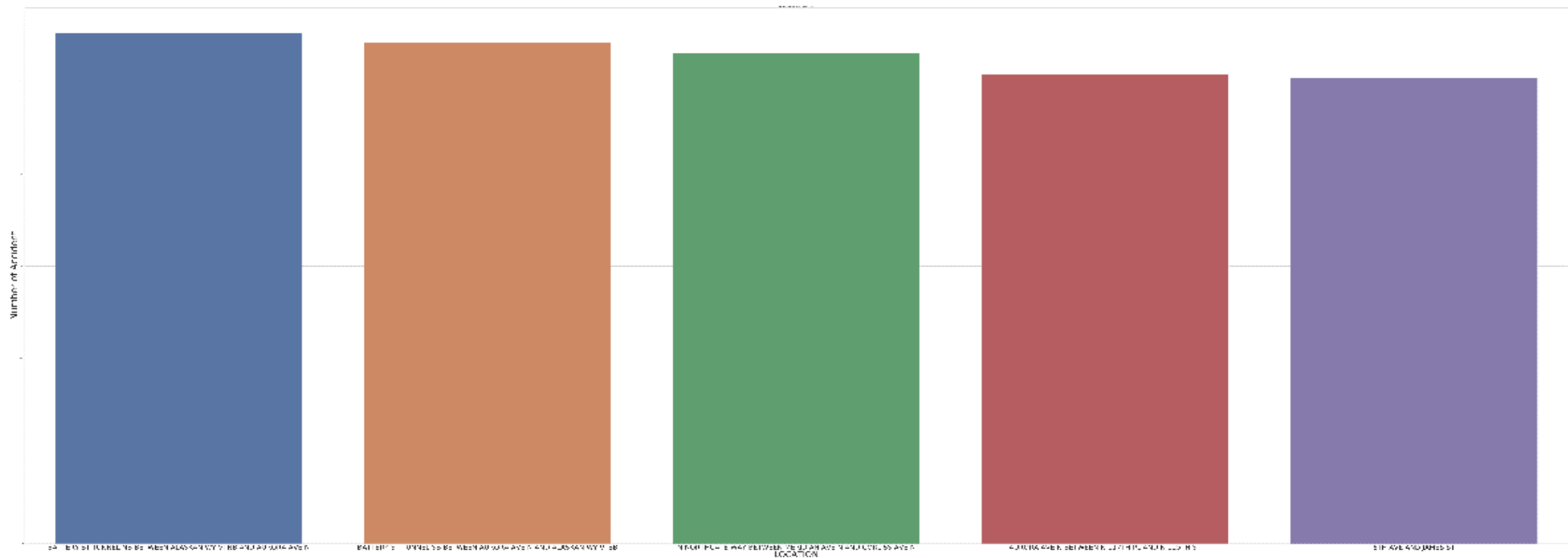
Weather
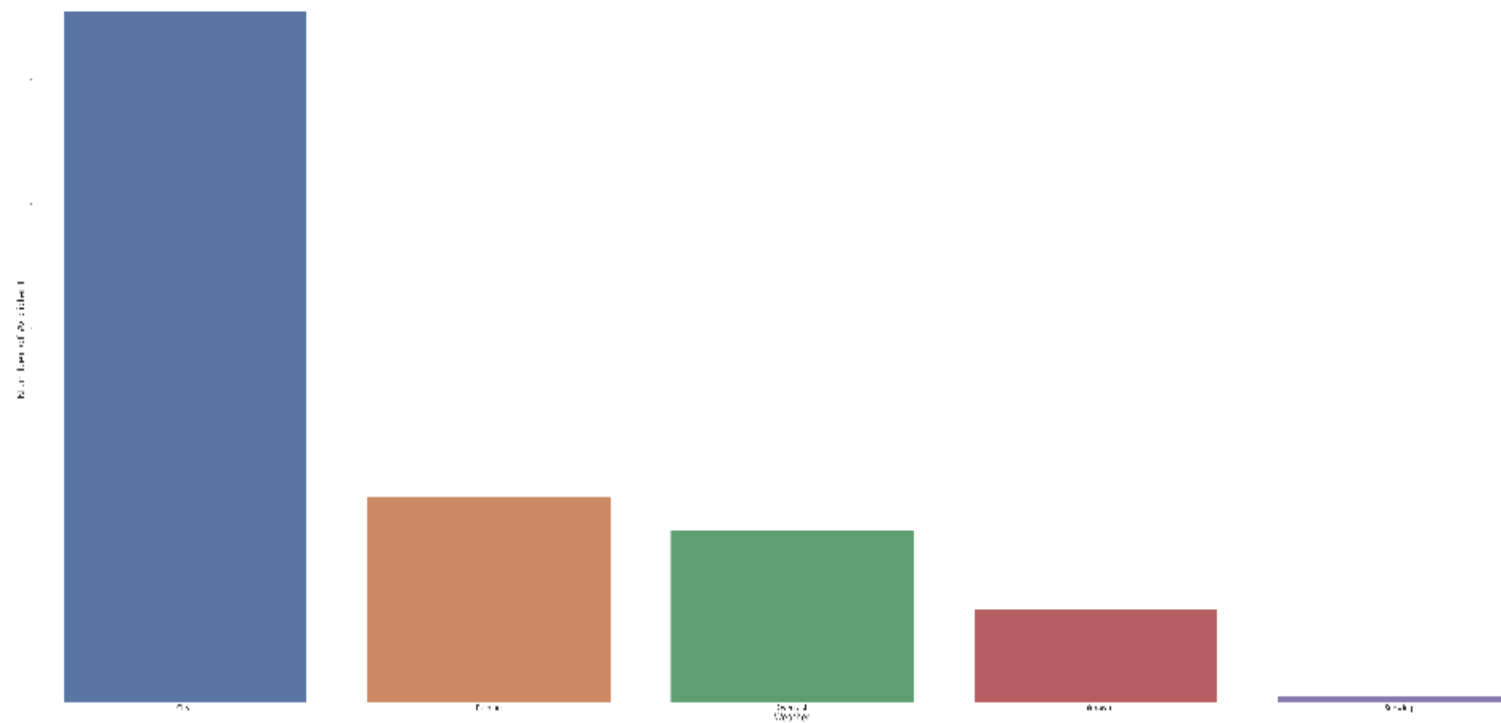
Road condition
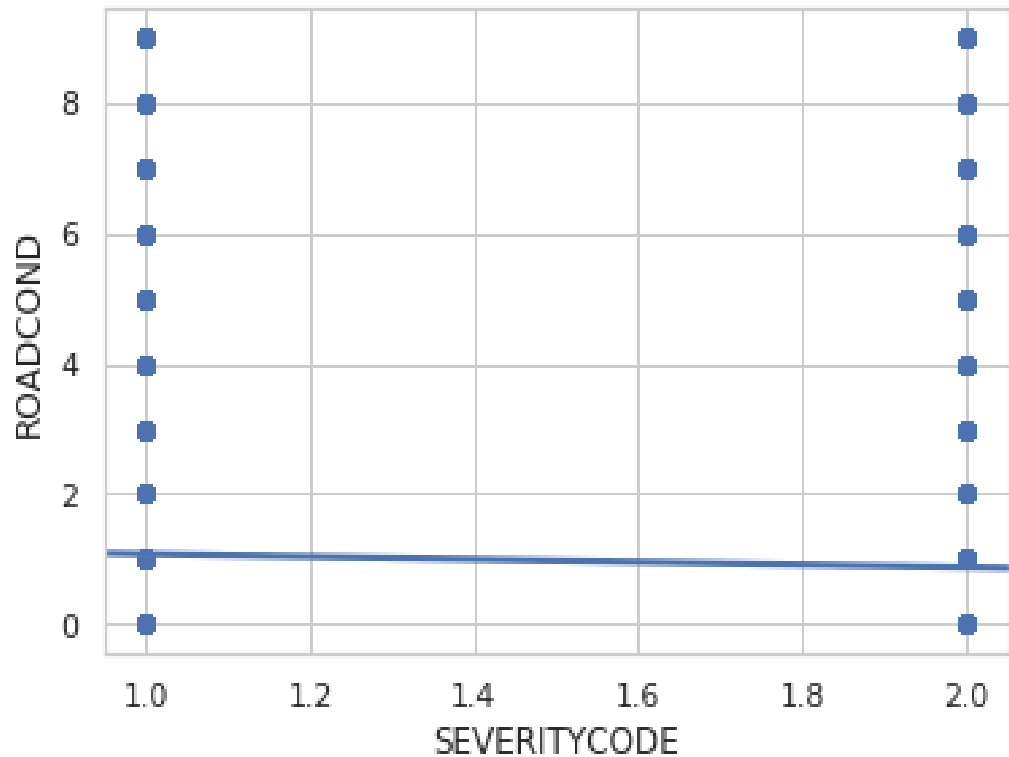
	ROADCOND	SEVERITYCODE
ROADCOND	1.000000	-0.082072
SEVERITYCODE	-0.082072	1.000000

Conclusion

- Clear Weather
- Dry road condition
- Speeding
- Intersection

Future direction

- Decision tree and classification model
- More datasets and variables
- Training and testing data