

City of Los Angeles: Traffic Collision Analysis

Study Group #2

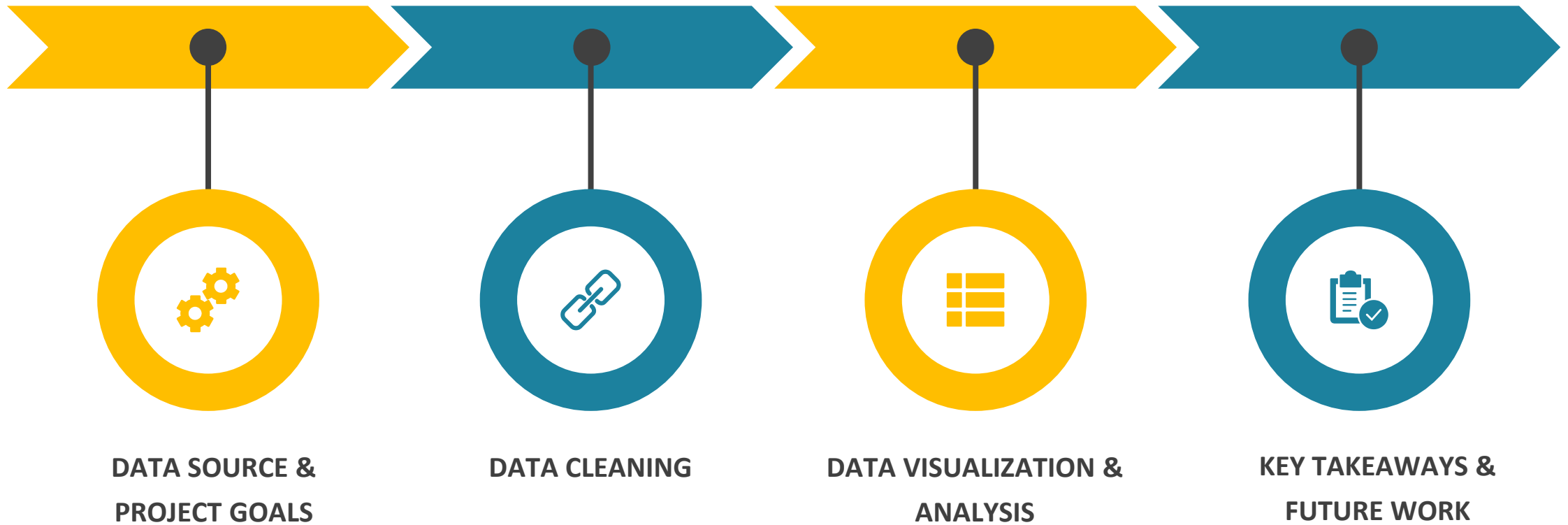
Hui Ye (Ines)

Veera Mouli Krishna Palacharla

Srikar Gunisetty

Xuefei Tan

OVERVIEW



Project Goals & Data Source

Project Data:

- City of Los Angeles Data Archives
- Real-time data dating back to 2010
- Exported in October 2021 in .csv format
- Data Size: 565,256 records x 18 features (~110 MB)
- Analysis tool: Python

[Click me for Data Source](#)

Goals:

- Exploratory Data Analysis
- Observe trends & patterns
- Help authorities (police departments and local DOTs) make data informed decisions to improve road safety
- Make recommendations on data collection

Data Cleaning

DATA



SORTED



ARRANGED

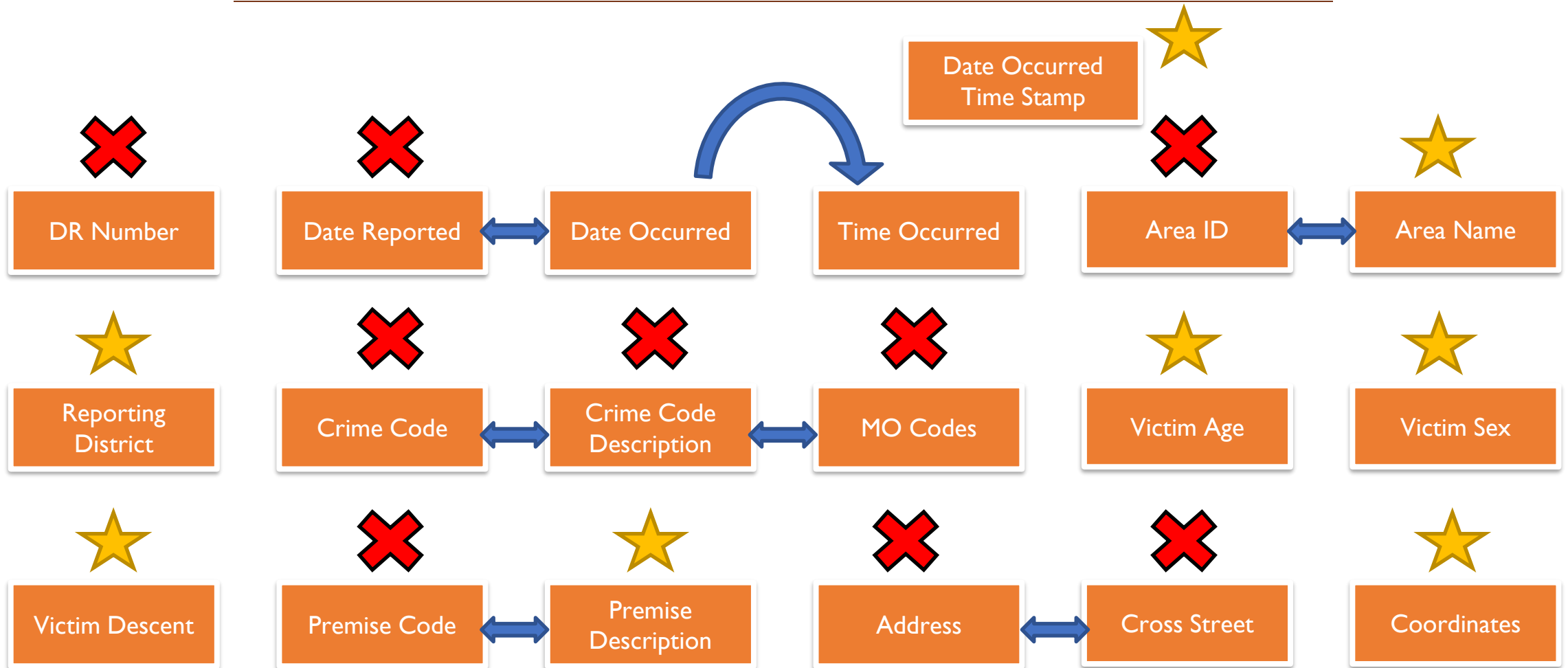


PRESENTED
VISUALLY

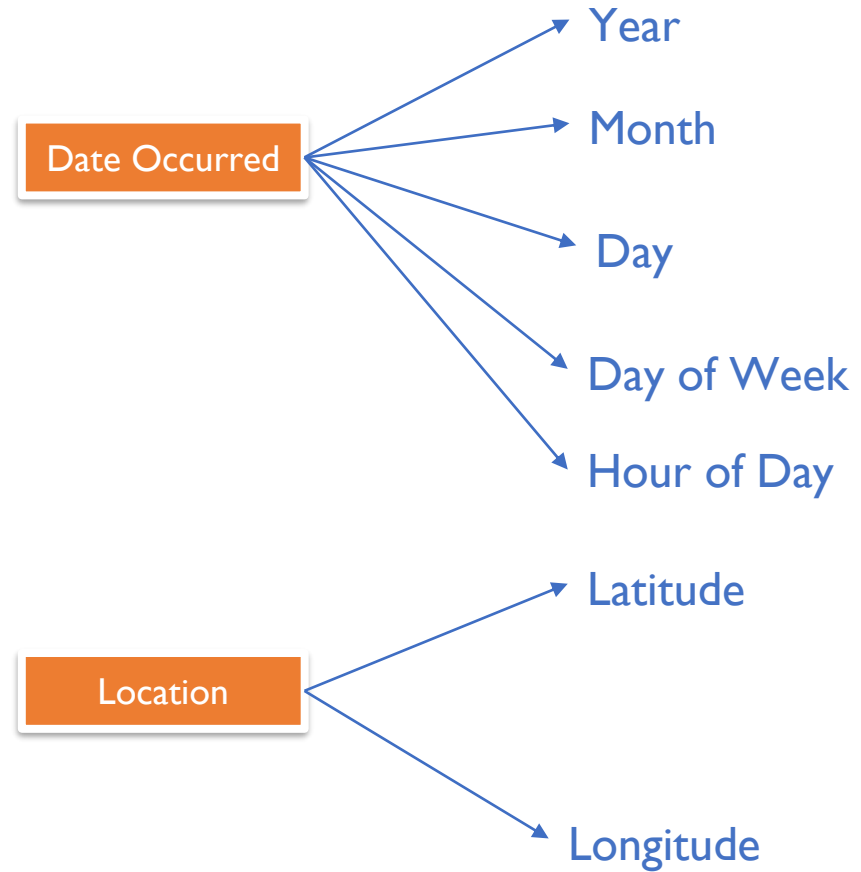


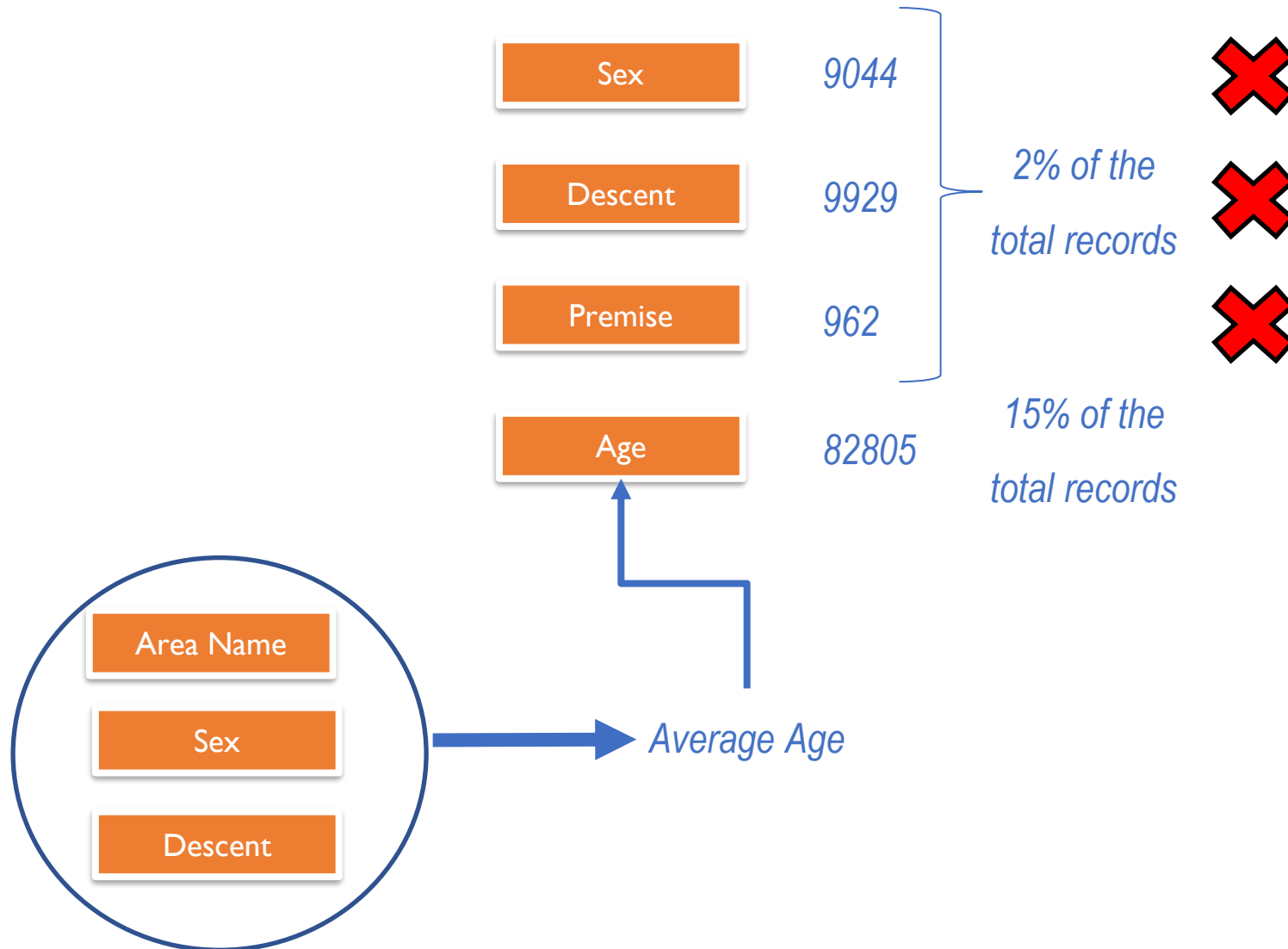
Tasks:

1. Identifying & discarding irrelevant data
2. Datatypes
3. Feature Engineering
4. Missing Values



Date Occurred	<i>object</i>	↔	<i>datetime64</i>
Area Name	<i>object</i>	↔	<i>category</i>
District	<i>object</i>	↔	<i>category</i>
Age	<i>int</i>	↔	<i>float</i>
Sex	<i>object</i>	↔	<i>category</i>
Descent	<i>object</i>	↔	<i>category</i>
Premise	<i>object</i>	↔	<i>category</i>
Location	<i>object</i>	↔	<i>category</i>

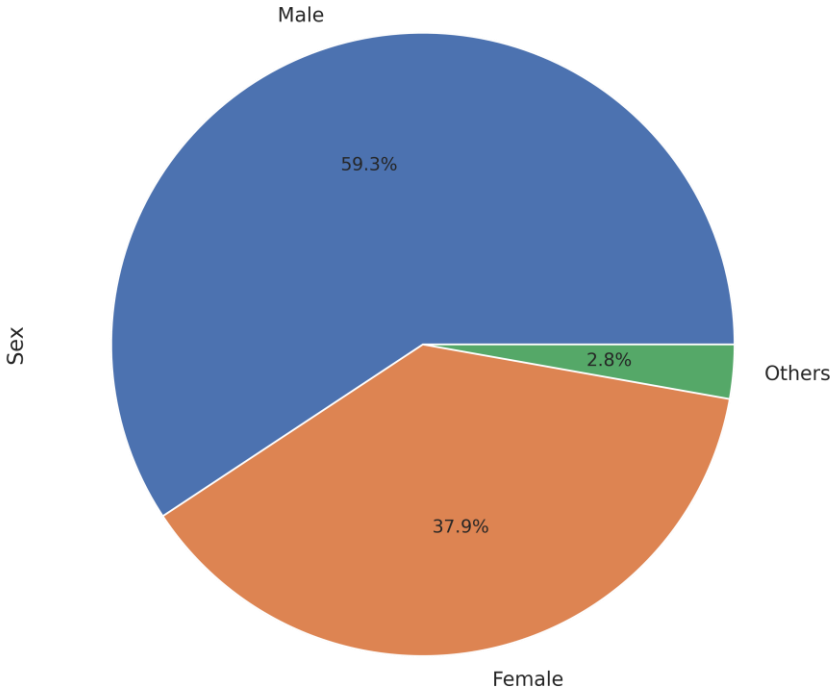




	Date_Occurred	Area_Name	District	Age	Sex	Descent	Premise	Year	Month	Day	Weekday	Latitude	Longitude	Weekday2	Date	Hour
0	2019-08-24 04:50:00	Southwest	356	22	M	H	STREET	2019	8	24	5	34.0255	-118.300	Saturday	2019-08-24	4
1	2019-08-30 23:20:00	Southwest	355	30	F	H	STREET	2019	8	30	4	34.0256	-118.308	Friday	2019-08-30	23
2	2019-08-25 05:45:00	Hollenbeck	422	37	M	X	STREET	2019	8	25	6	34.0738	-118.207	Sunday	2019-08-25	5
3	2019-11-20 03:50:00	Central	128	21	M	H	STREET	2019	11	20	2	34.0492	-118.239	Wednesday	2019-11-20	3
4	2019-08-30 21:00:00	Southwest	374	49	M	B	STREET	2019	8	30	4	34.0108	-118.318	Friday	2019-08-30	21

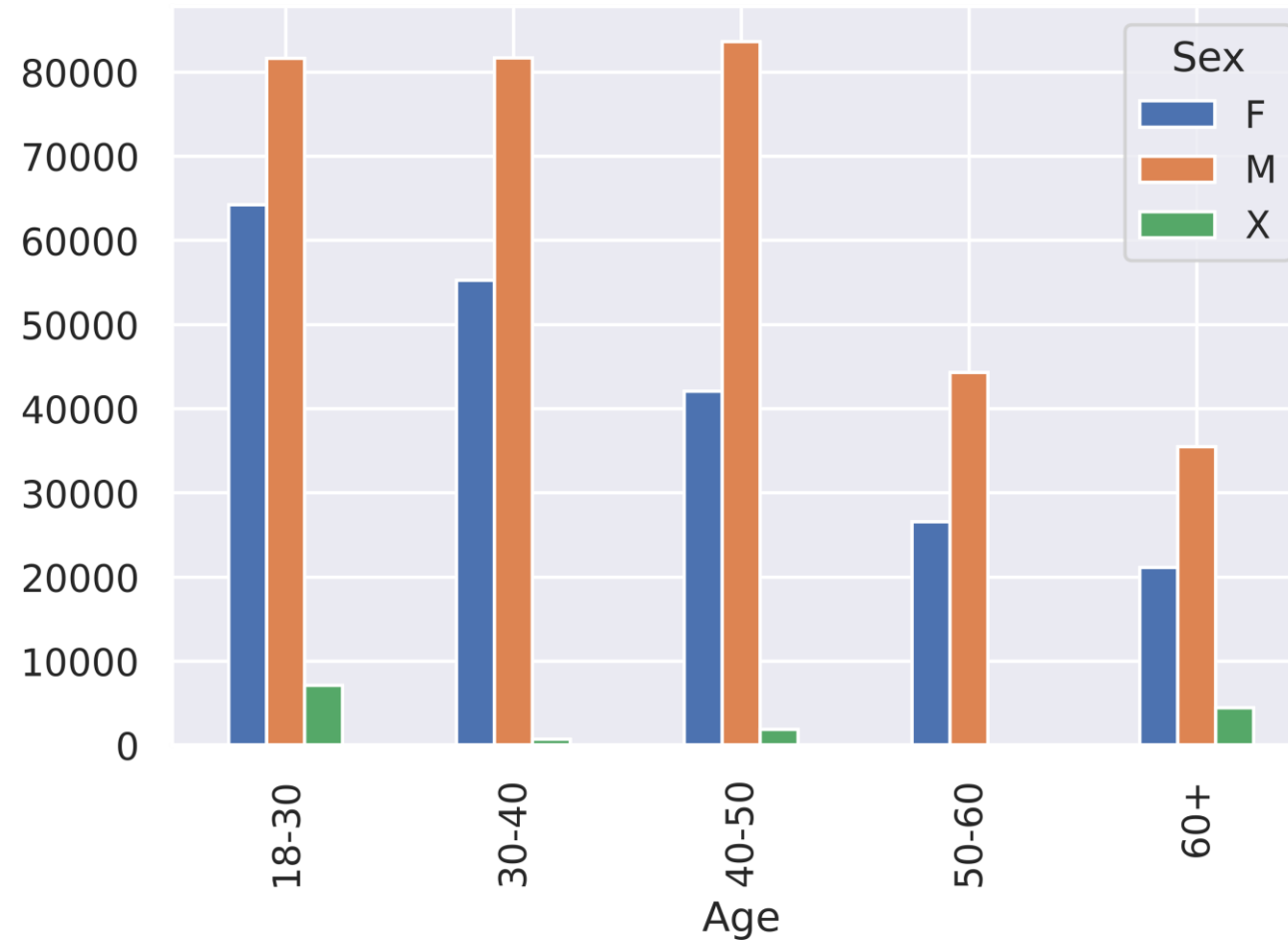
Data Visualizations and Analysis

Distribution of Collisions with Gender

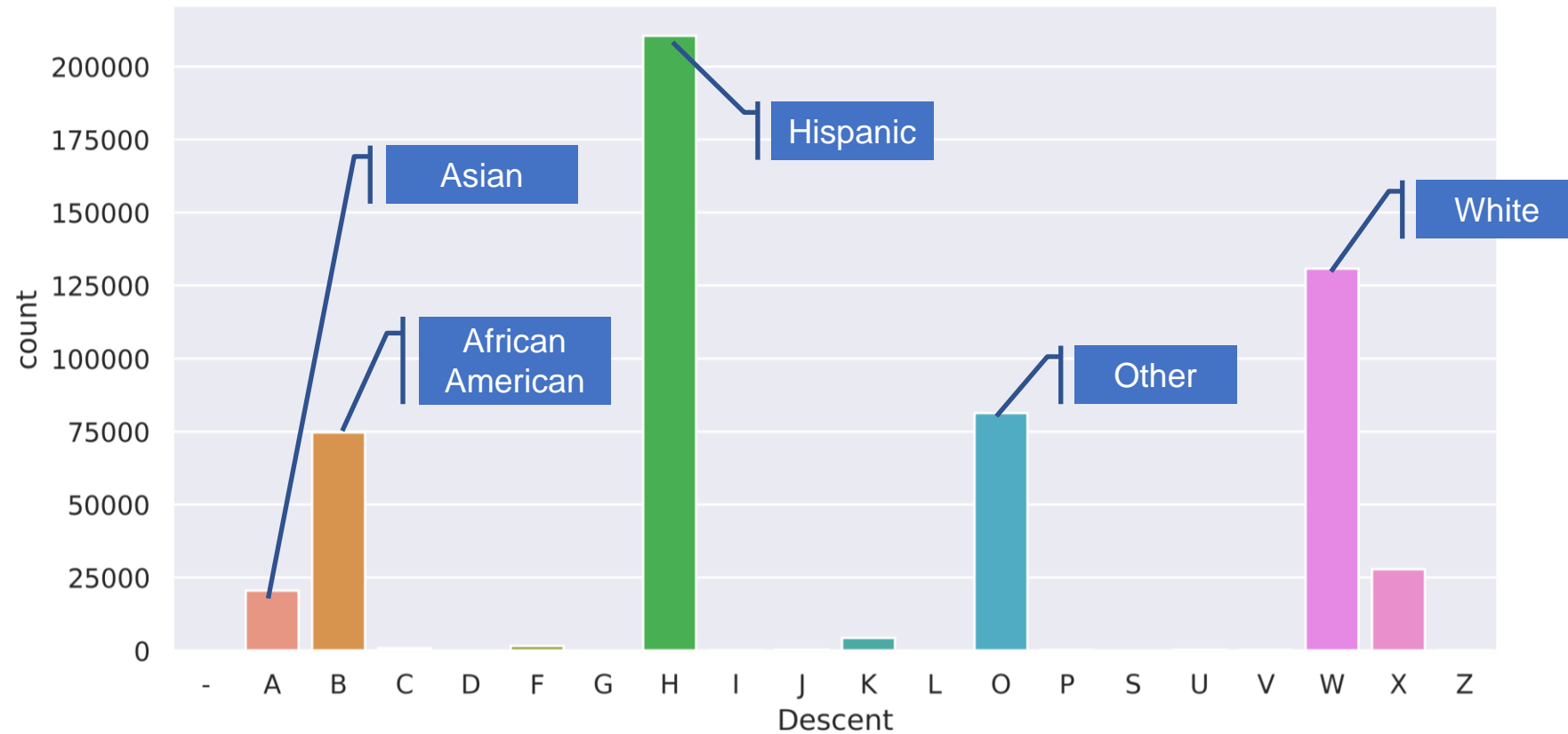


Gender	Average Age
Female	40
Male	42
Other	39

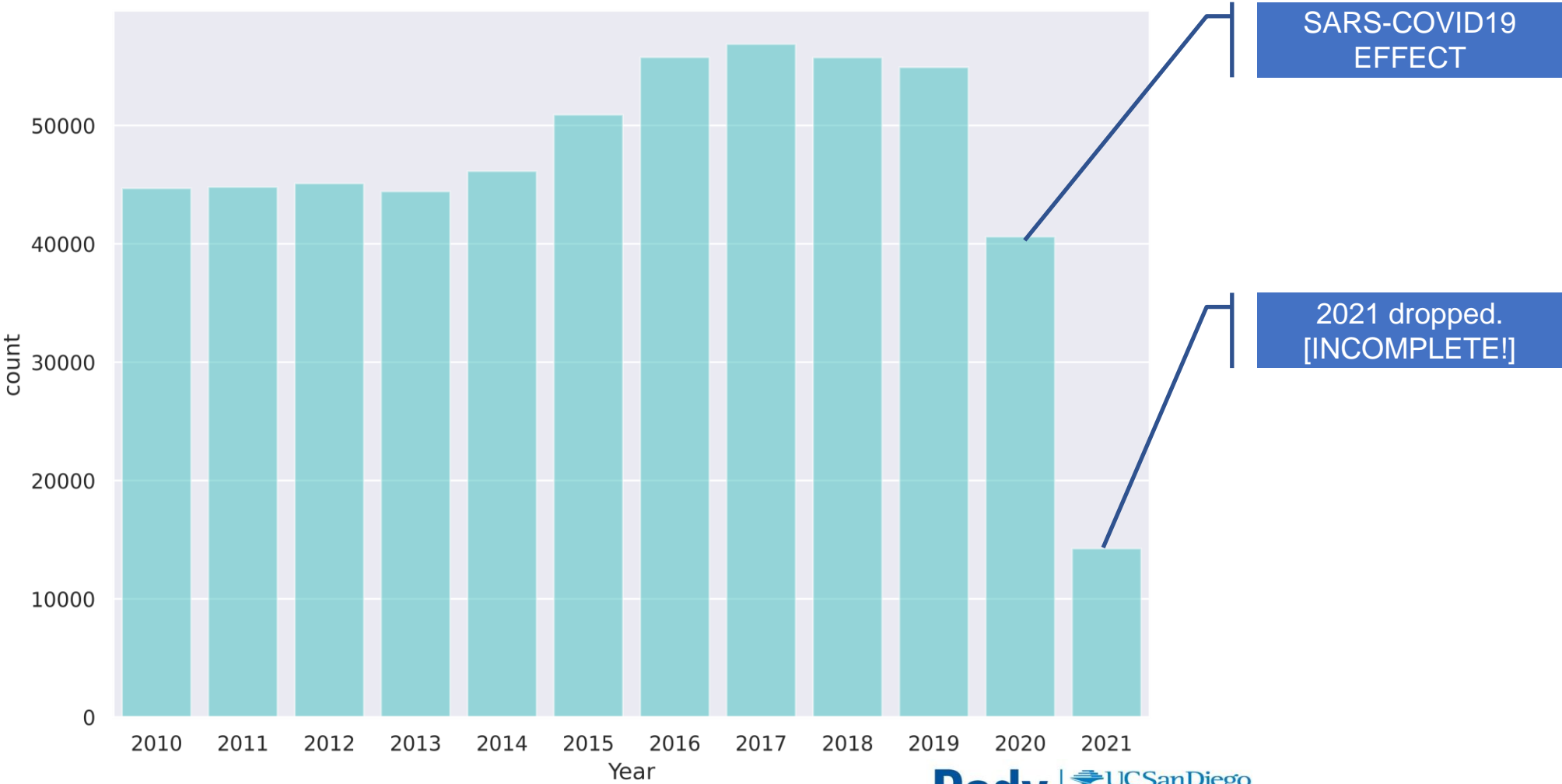
Distribution of Collisions w/ Age



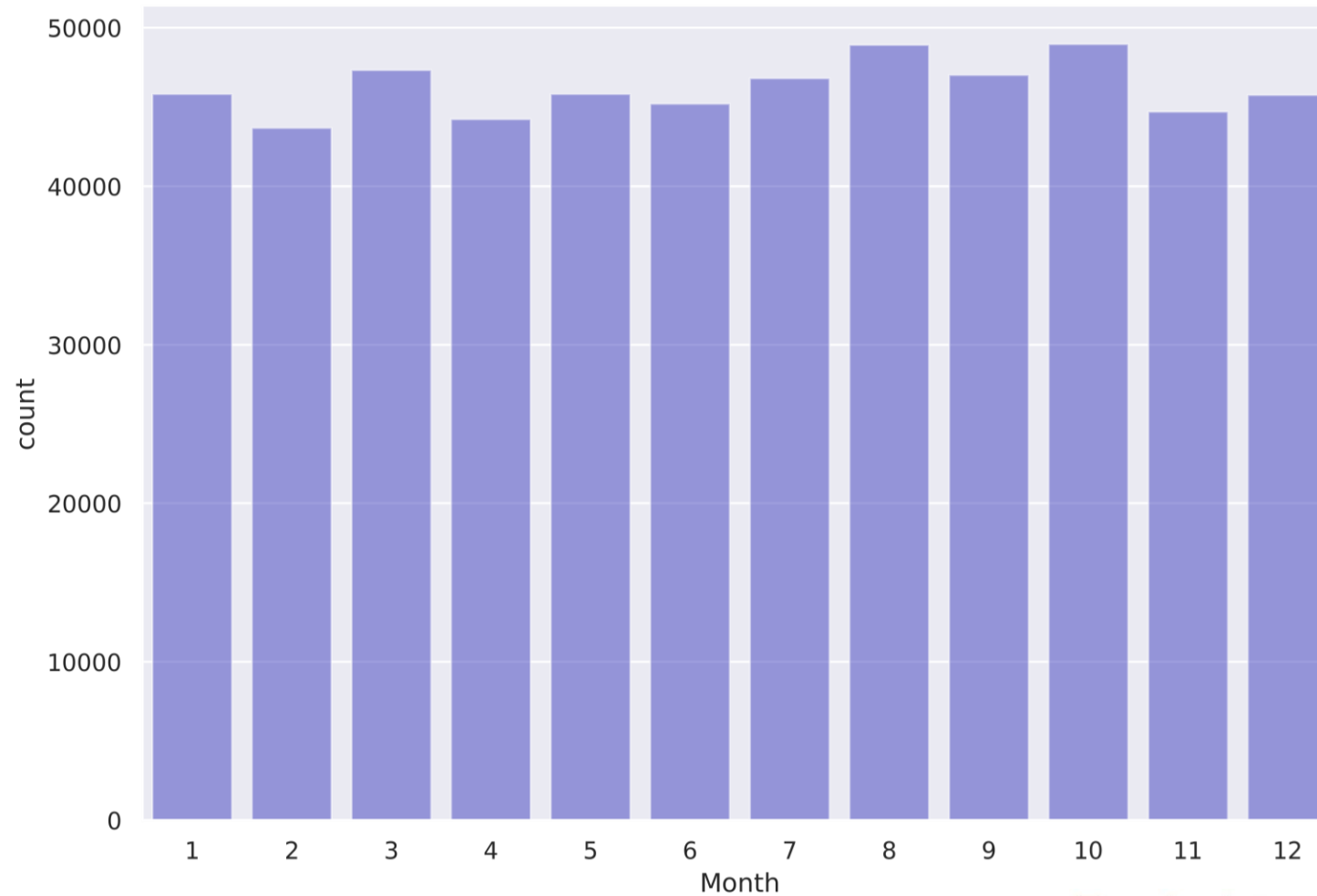
Distributions of Collisions by Descent



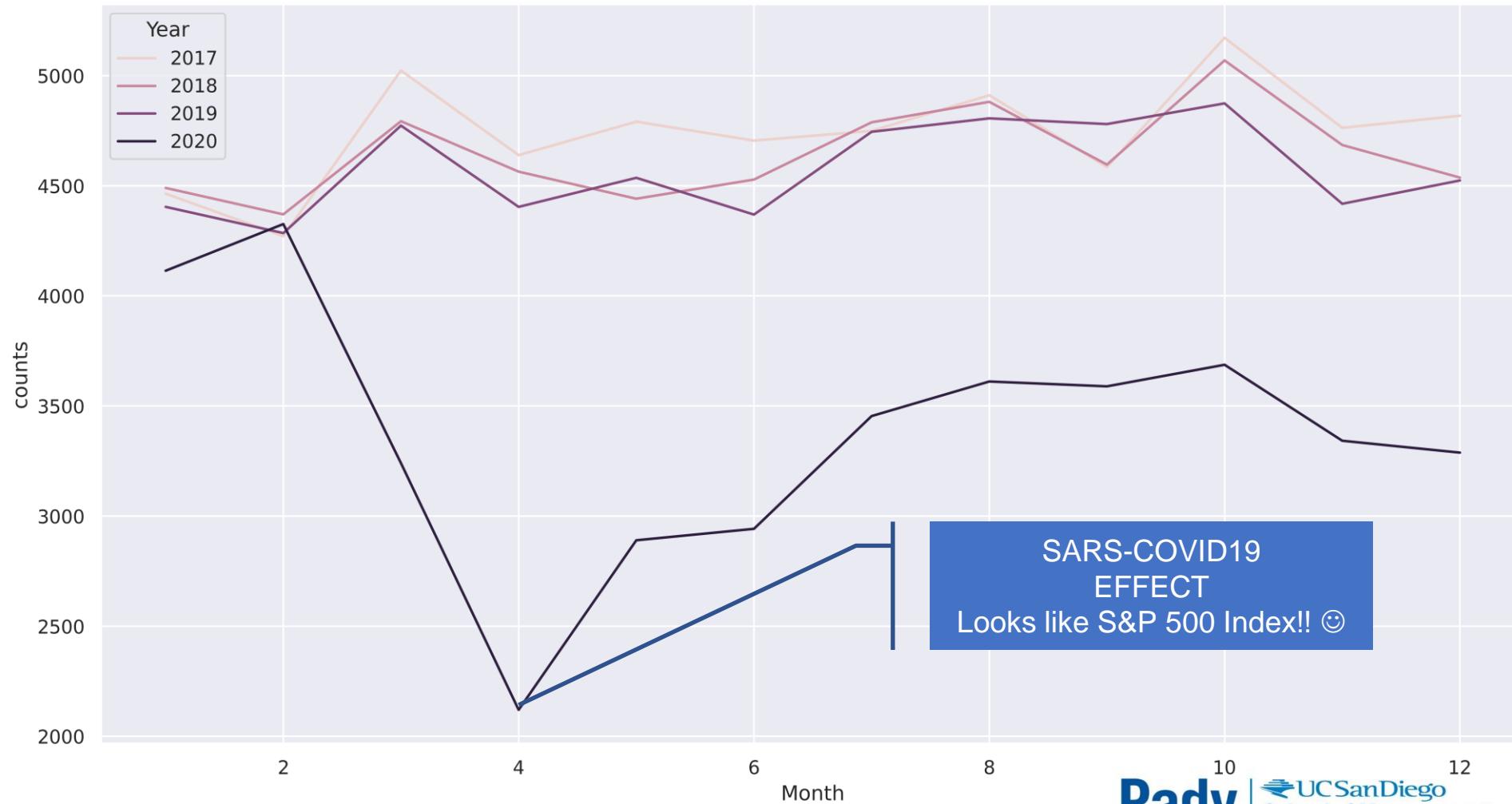
Distribution of Collisions by Year



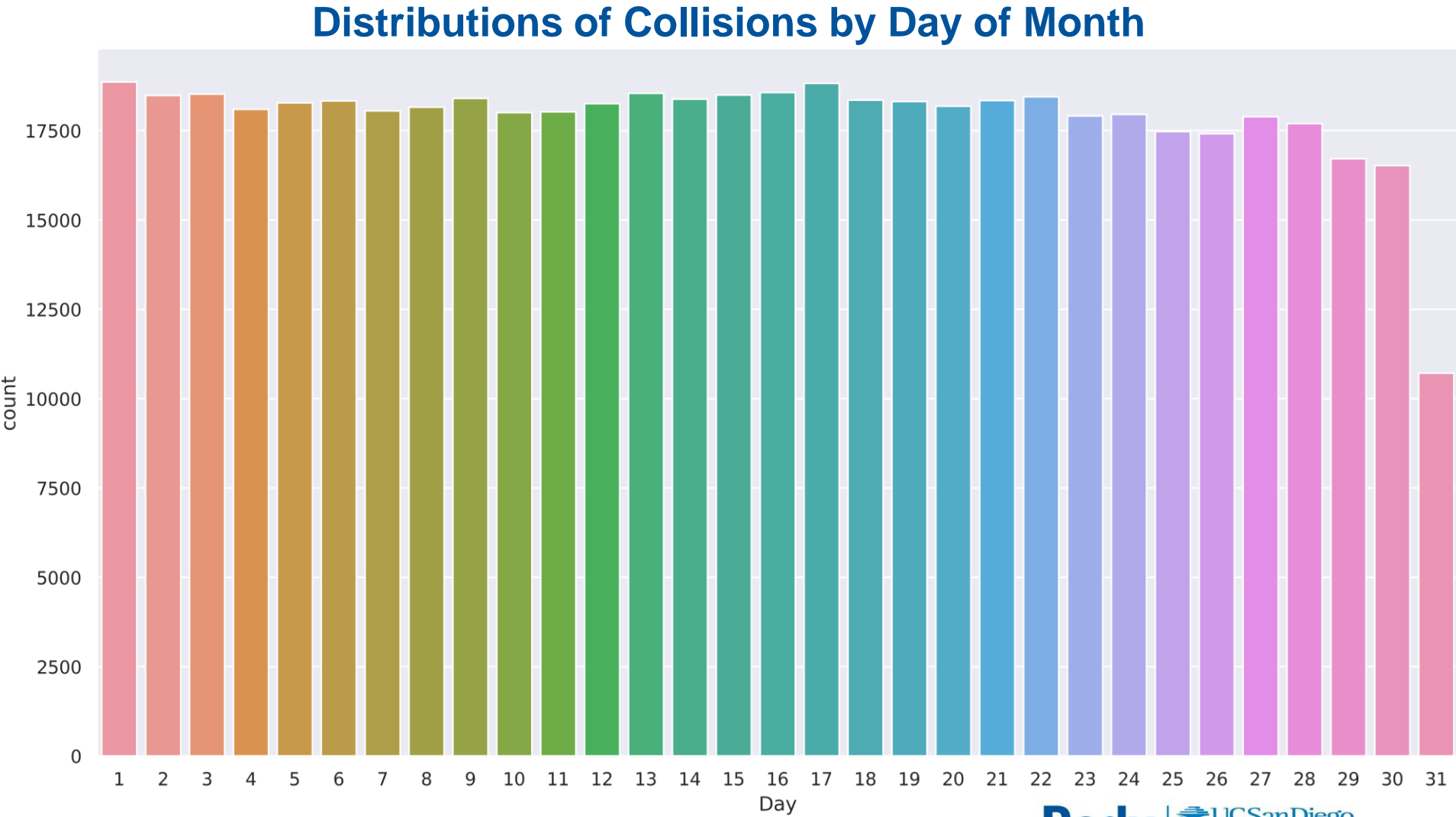
Distribution of Collisions by Month



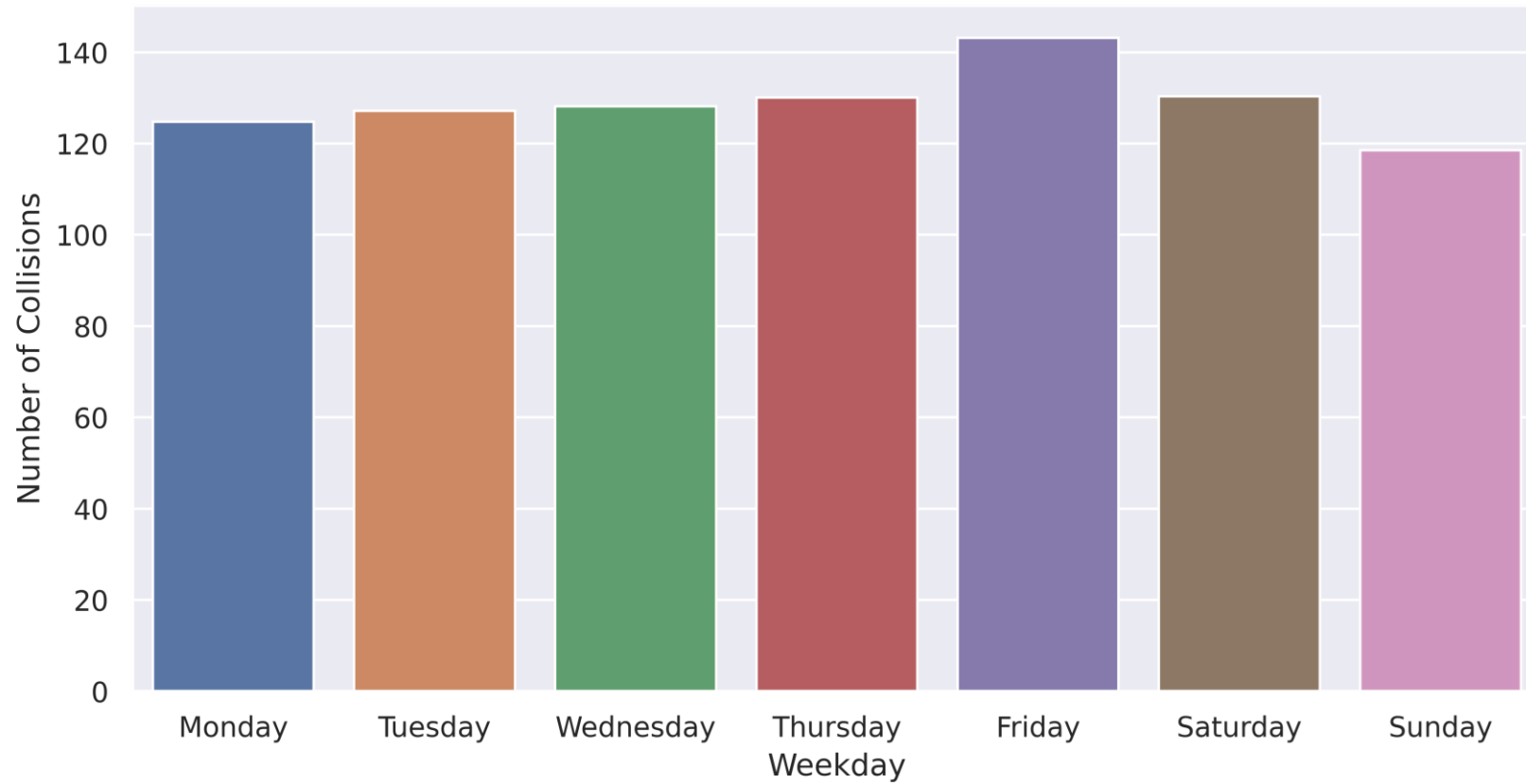
Collisions by Month from 2017-2020



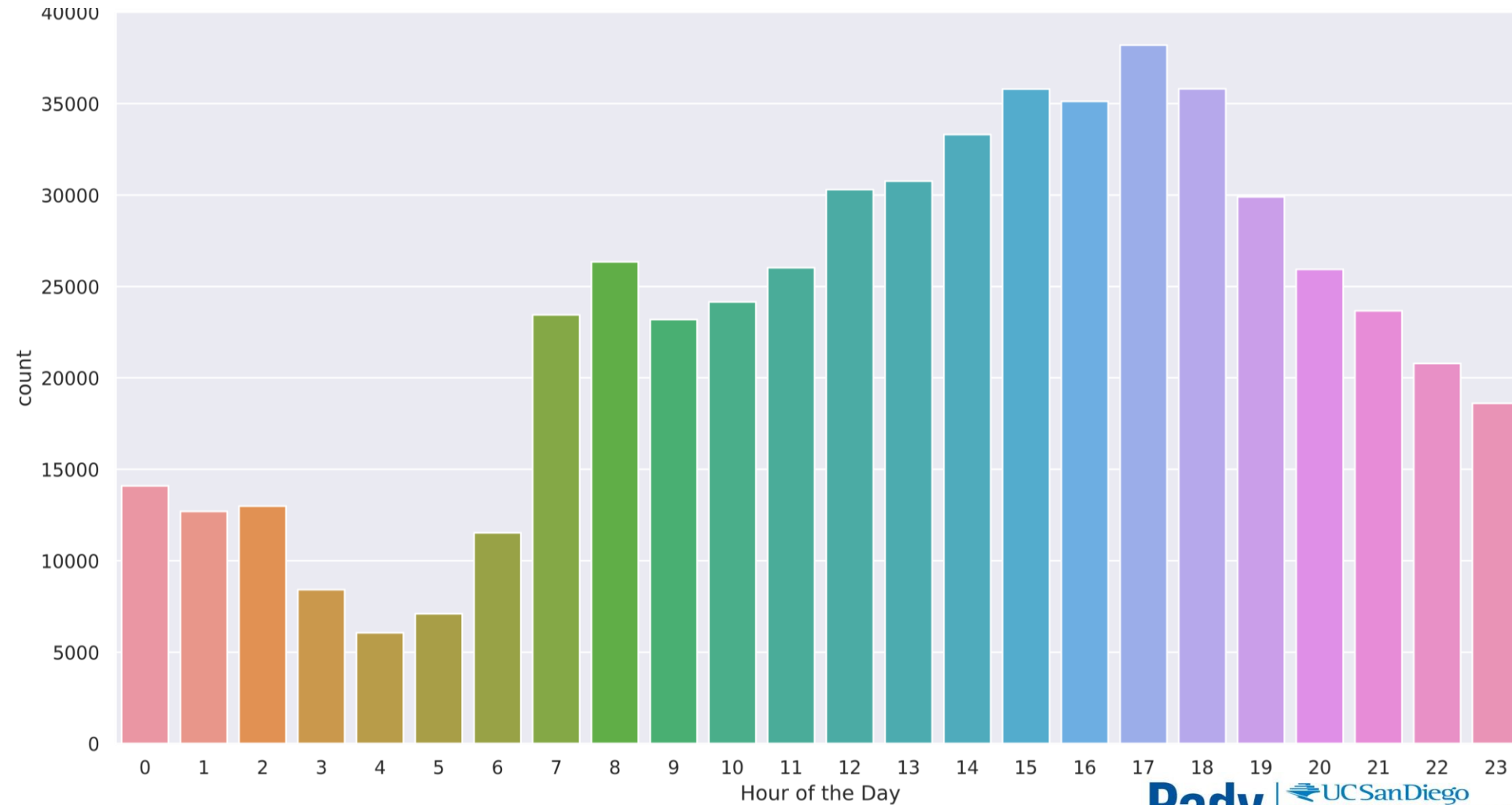
SARS-COVID19
EFFECT
Looks like S&P 500 Index!! 😊



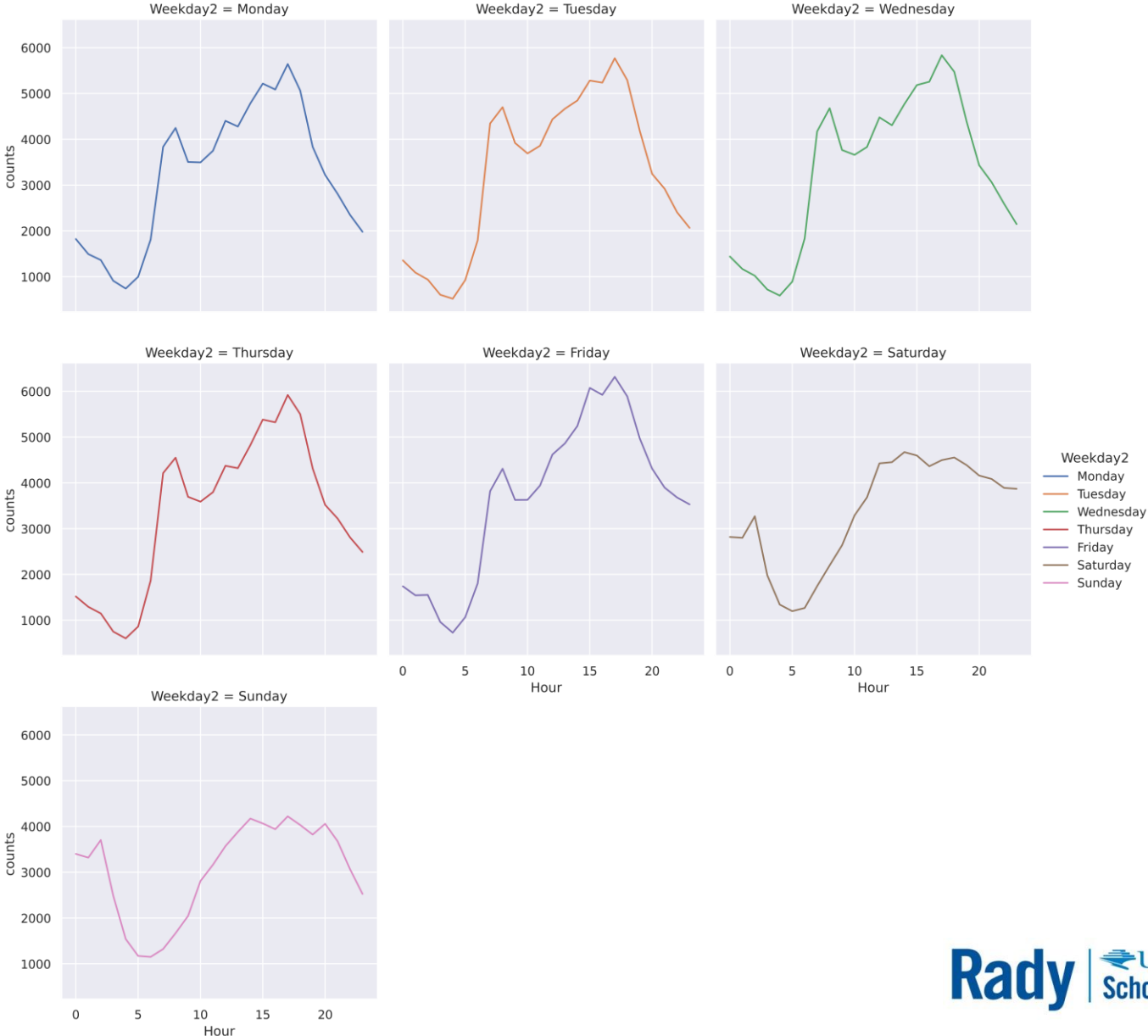
Distributions of Average # of Collisions by Weekday



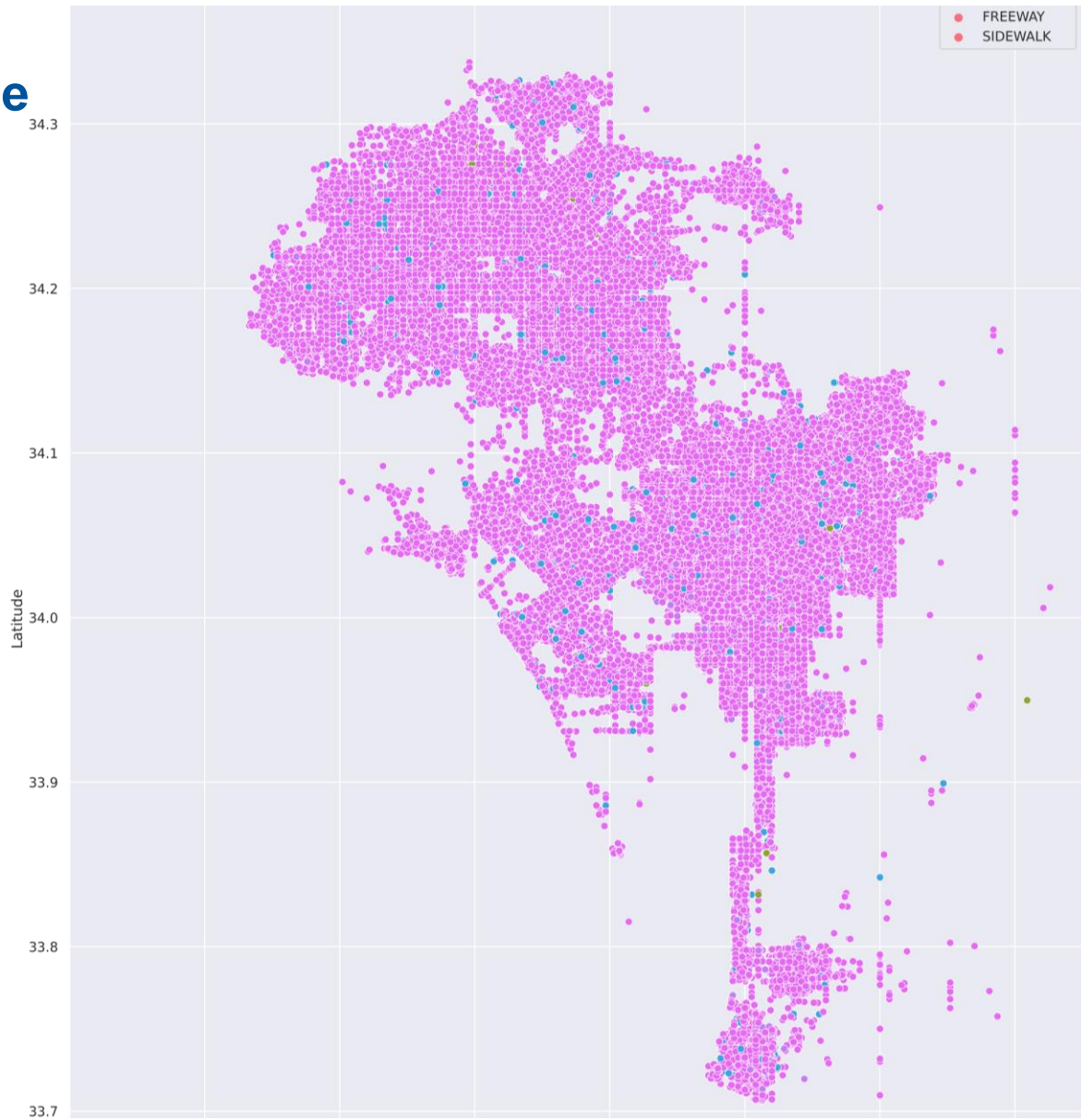
Distributions of Collisions by Hour of day

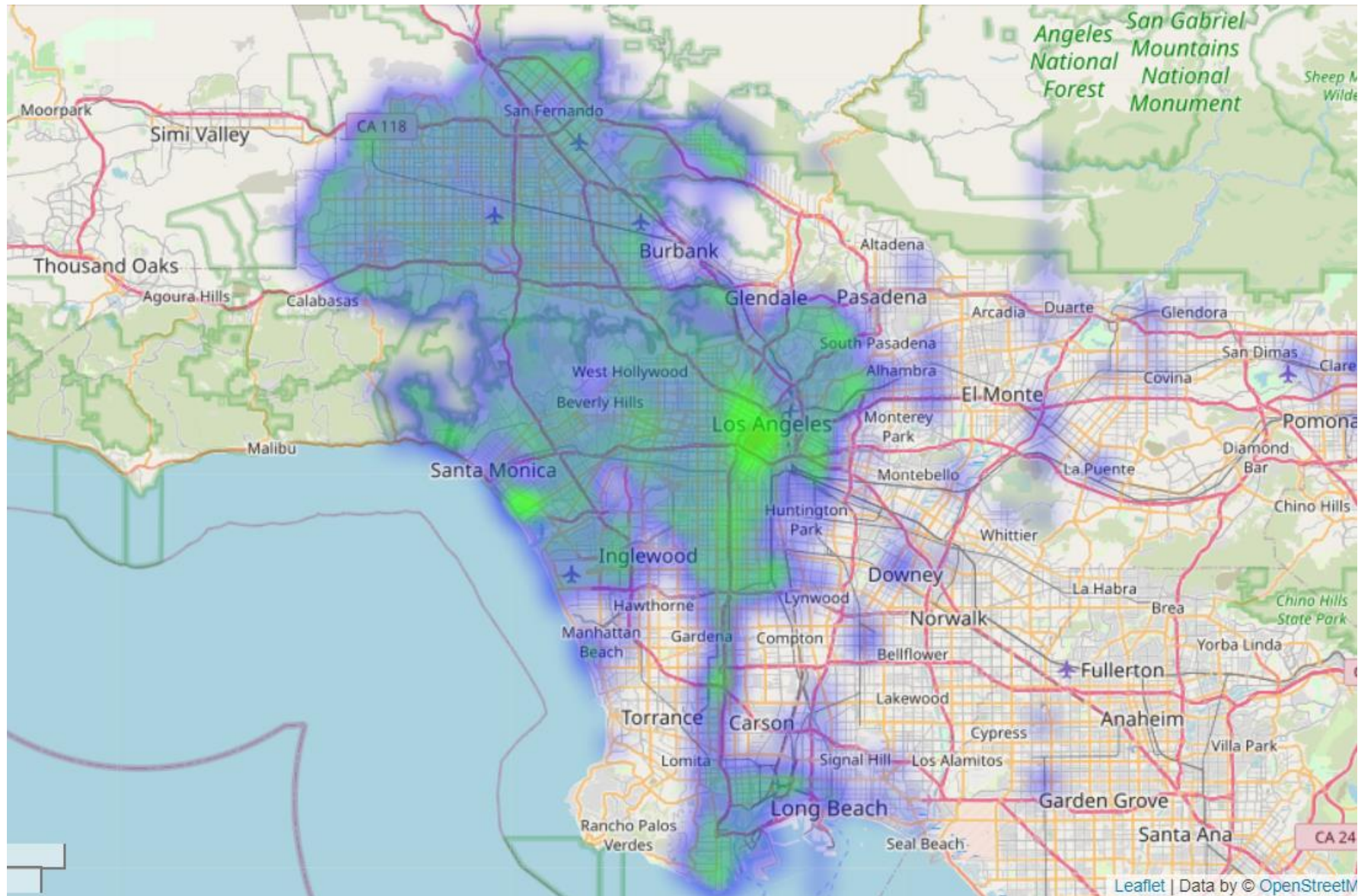


Distributions of Collisions by Hour & Day of Week

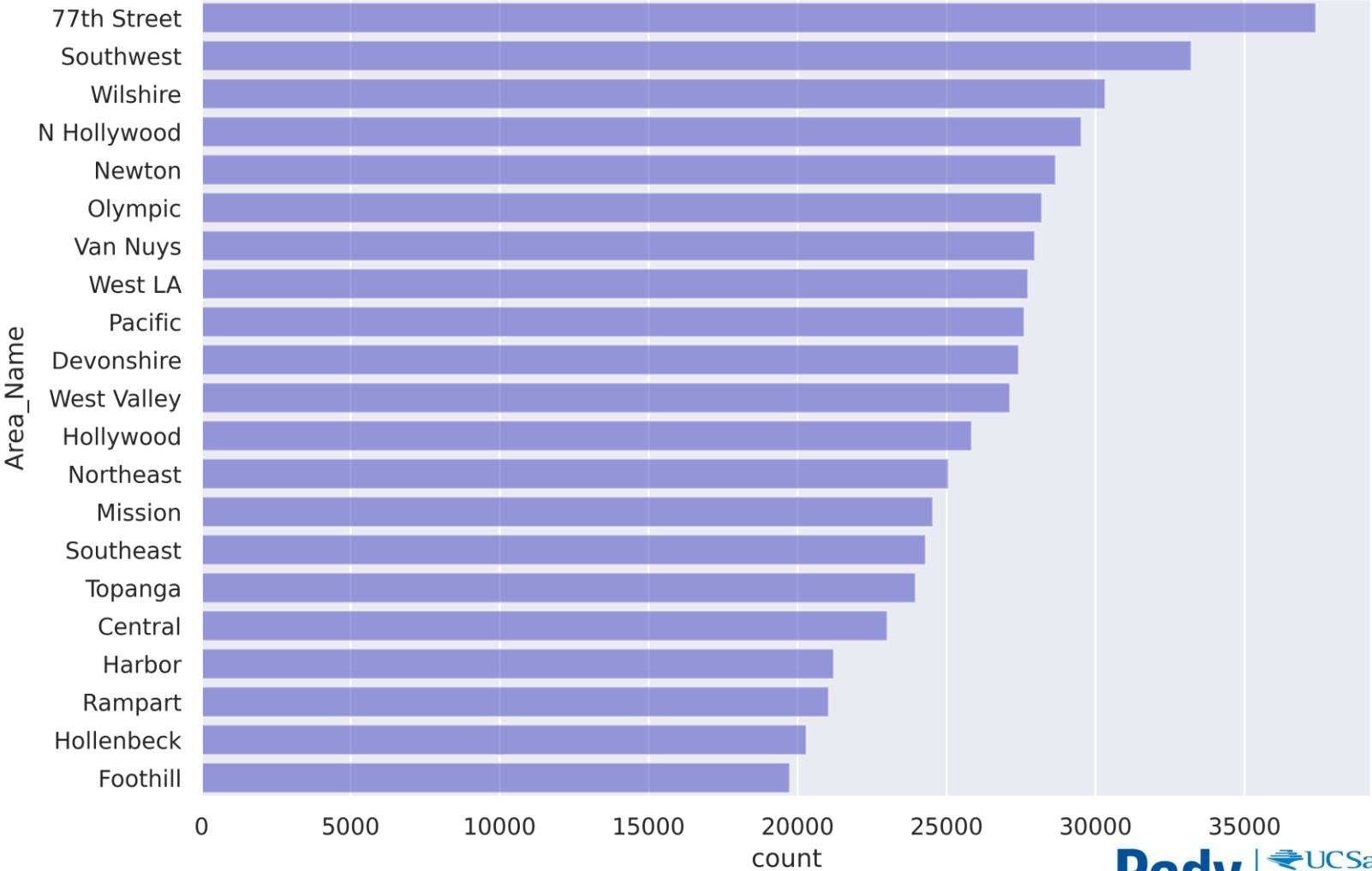


Collisions by Premise





Collisions by Area Name



Key Takeaways:

1. Overall decrease in the collisions from 2010
2. Identified days, time frames for busiest & safest streets
3. Identified critical neighborhoods to help local authorities implement additional traffic resources such as higher patrols, safety signs, roadway improvements etc.
4. Break stereotypes

Future Work:

1. Expand framework to multiple cities
2. Predictive Analytics (probability of collision)
3. Data Collection recommendations
 - Vehicle type
 - Scale of injury
 - Cause of collision
 - Weather conditions



Thank You 😊

Questions?
