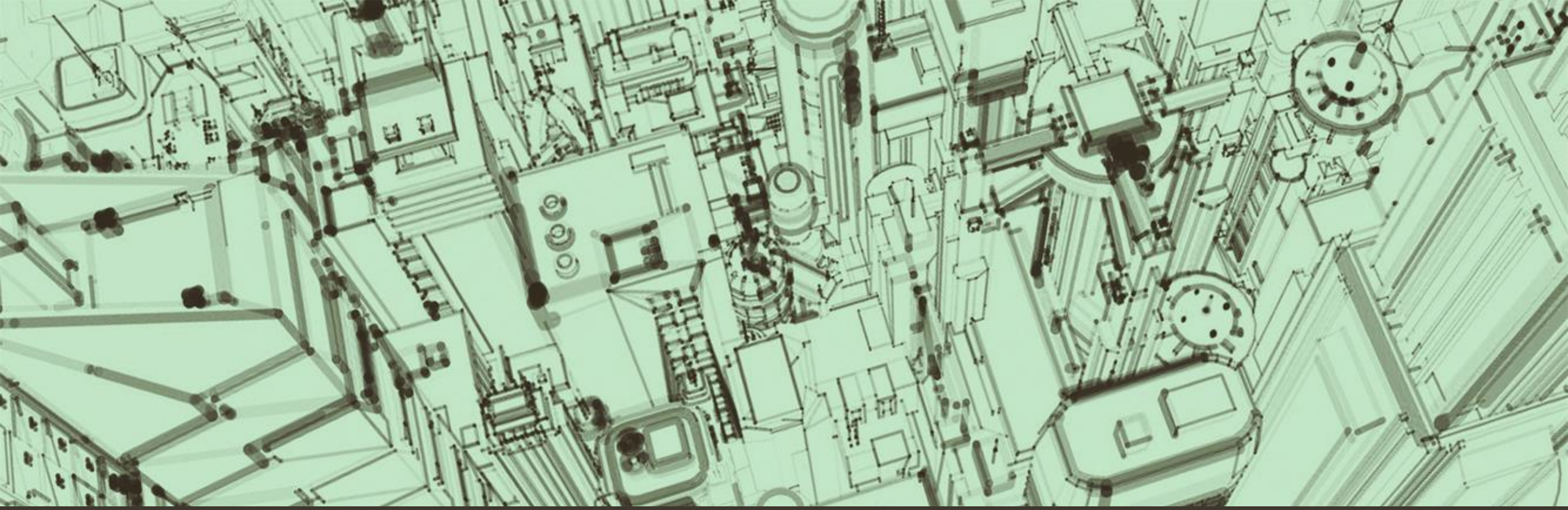


Washington Fatal Crash Files

Team 026 - Jimmy Nguyen, John Dutan, Anthony Capasso

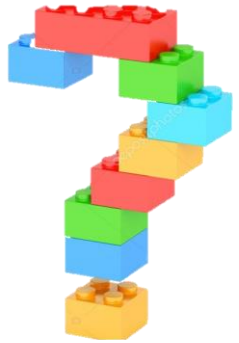
Table of Contents

Overview

Overview

Washington State's **Official Highway Safety Office**. Manages traffic safety programs in **media and outreach**.

Plan to reduce traffic deaths or injuries on roadways to **zero** by the year 2030.



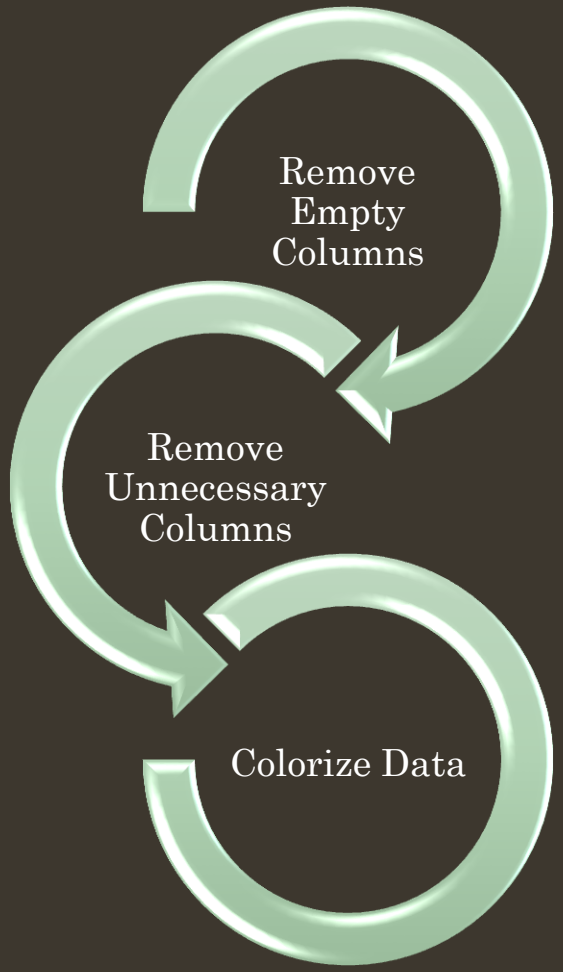
Are the people who **live in these communities** the same people involved in **fatal crashes that occur there**?

Overview

Thought-Process

Data Clean

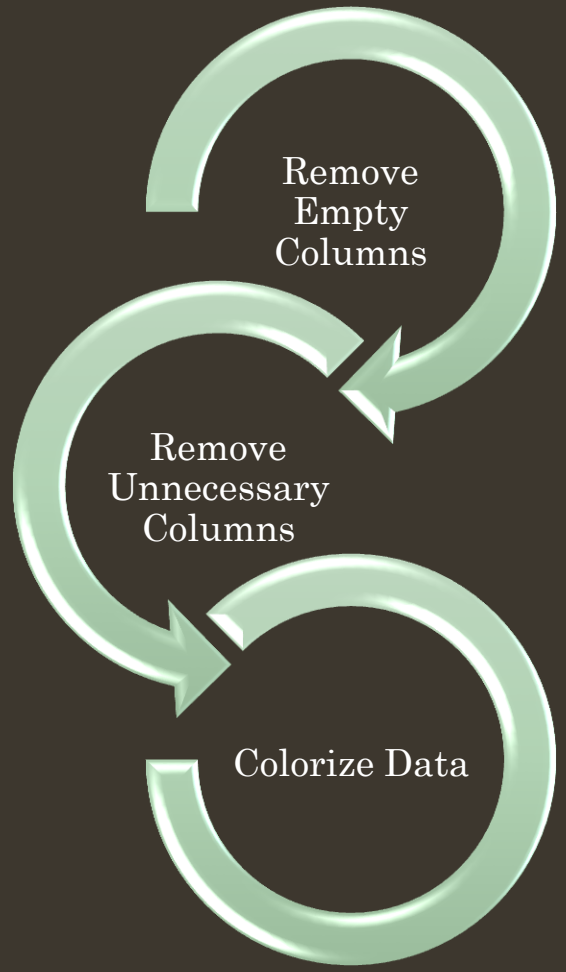
Thought-Process



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	year	case	par	repjur	crash_dt	crash_tm	accday	accmon	holiday	county	co_char	city	weather	weather2	lightcond	surfcond	surftype	numfatal
2	2017	1	E628946	2	1/1/2017	2:12	1	1	1	61	Snohomish	0	3	0	2	4	2	1
3	2017	2	E627989	26	1/2/2017	17:14	2	1	1	35	Kitsap	0	1	0	2	1	2	1
4	2017	4	3747633	263	1/1/2017	18:47	1	1	1	33	King	1960	1	0	3	1	1	1
5	2017	5	E628691	4	1/1/2017	3:50	1	1	1	63	Spokane	2110	4	0	3	3	2	1
6	2017	6	3746306	263	1/5/2017	9:53	5	1	0	33	King	1960	1	0	1	1	2	1
7	2017	6	3746306	263	1/5/2017	9:53	5	1	0	33	King	1960	1	0	1	1	2	1
8	2017	6	3746306	263	1/5/2017	9:53	5	1	0	33	King	1960	1	0	1	1	2	1
9	2017	7	E629178	5	1/6/2017	18:48	6	1	0	41	Lewis	300	1	0	2	1	2	1
10	2017	8	E629981	6	1/7/2017	14:12	7	1	0	37	Kittitas	0	4	0	1	3	1	1
11	2017	8	E629981	6	1/7/2017	14:12	7	1	0	37	Kittitas	0	4	0	1	3	1	1
12	2017	8	E629981	6	1/7/2017	14:12	7	1	0	37	Kittitas	0	4	0	1	3	1	1
13	2017	9	E632204	4	1/8/2017	16:39	8	1	0	1	Adams	0	3	0	2	4	2	1
14	2017	9	E632204	4	1/8/2017	16:39	8	1	0	1	Adams	0	3	0	2	4	2	1
15	2017	10	E632128	8	1/8/2017	19:05	8	1	0	45	Mason	0	2	0	2	2	2	1
16	2017	10	E632128	8	1/8/2017	19:05	8	1	0	45	Mason	0	2	0	2	2	2	1
17	2017	11	E630209	40	1/7/2017	9:39	7	1	0	63	Spokane	0	1	0	1	4	2	2
18	2017	12	E631317	7	1/11/2017	2:05	11	1	0	33	King	1960	1	0	3	1	1	1
19	2017	12	E631317	7	1/11/2017	2:05	11	1	0	33	King	1960	1	0	3	1	1	1
20	2017	13	E630362	42	1/9/2017	6:51	9	1	0	67	Thurston	0	1	0	2	4	2	1
21	2017	14	E631092	83	1/6/2017	6:01	6	1	0	41	Lewis	310	1	0	2	1	2	1
22	2017	15	E631425	35	1/11/2017	20:23	11	1	0	53	Pierce	0	1	0	3	1	2	1
23	2017	16	E632347	6	1/14/2017	3:40	14	1	2	37	Kittitas	0	1	0	2	1	2	1
24	2017	16	E632347	6	1/14/2017	3:40	14	1	2	37	Kittitas	0	1	0	2	1	2	1
25	2017	17	E634438	3	1/15/2017	14:21	15	1	2	53	Pierce	1124	1	0	1	1	2	1
26	2017	18	E631960	2	1/13/2017	18:40	13	1	2	73	Whatcom	150	1	0	3	1	2	1
27	2017	18	E631960	2	1/13/2017	18:40	13	1	2	73	Whatcom	150	1	0	3	1	2	1
28	2017	18	E631960	2	1/13/2017	18:40	13	1	2	73	Whatcom	150	1	0	3	1	2	1
29	2017	19	E634659	7	1/13/2017	21:06	13	1	2	33	King	2310	1	0	3	1	1	1
30	2017	19	E634659	7	1/13/2017	21:06	13	1	2	33	King	2310	1	0	3	1	1	1
31	2017	20	E633291	184	1/17/2017	21:05	17	1	0	61	Snohomish	1215	2	0	3	2	2	1
32	2017	21	E634199	8	1/18/2017	0:09	18	1	0	35	Kitsap	0	2	0	2	2	2	1
33	2017	22	E635015	1	1/19/2017	7:30	19	1	0	77	Yakima	0	1	0	1	3	1	1
34	2017	22	E635015	1	1/19/2017	7:30	19	1	0	77	Yakima	0	1	0	1	3	1	1
35	2017	22	E635015	1	1/19/2017	7:30	19	1	0	77	Yakima	0	1	0	1	3	1	1

Data Clean

Thought-Process

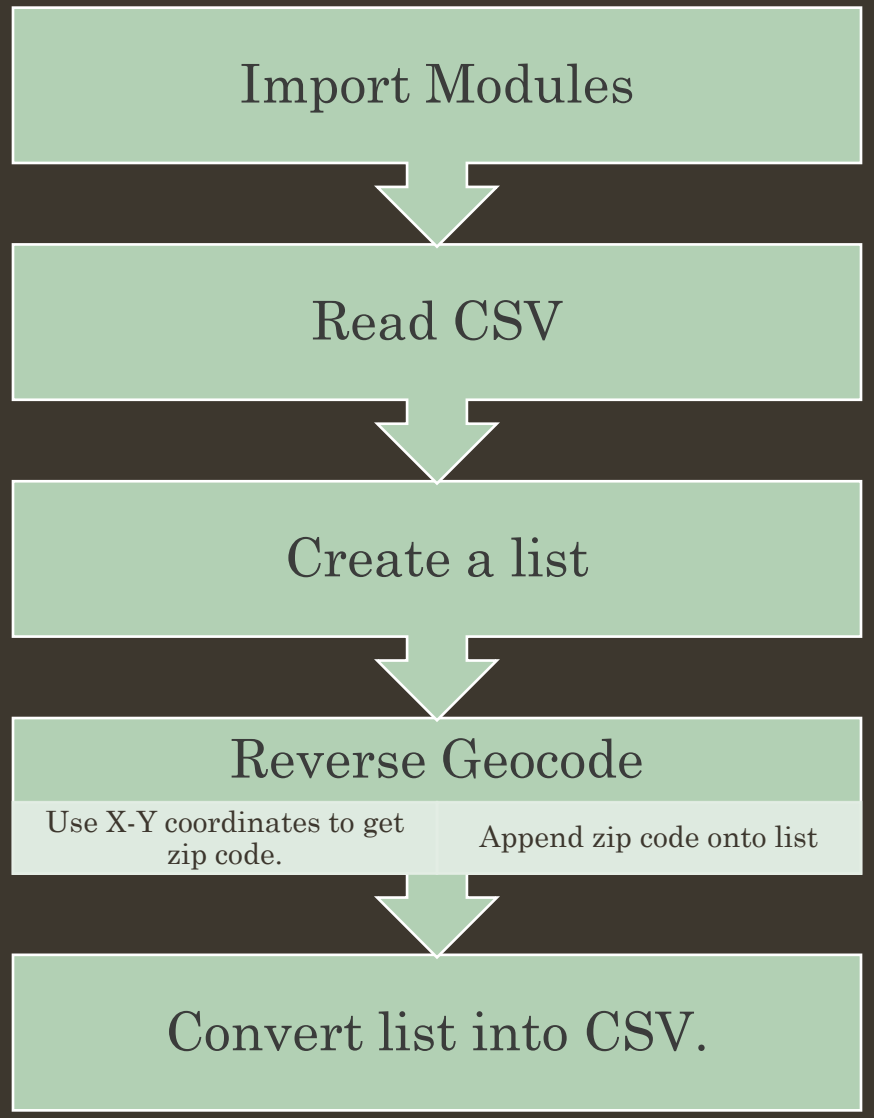


surfcond	surftype	trafid1	trafid2	x	y	regowner	regstate	seqevent1	seqevent2	seqevent3	seqevent4	seqevent5	seqevent6	age	sex	hispanic	race
4	2	I-5		-122.1756167	48.02354722	2	53	63	1	3				26	1	1	
1	2	CR-SE BURLEY	OLALLA RD	-122.583225	47.43134167	1	53	63	69	68	64	68	63	74	1	7	
1	1	WALLINGFORD	N 82ND ST	-122.3364222	47.6887	2	53	8						27	2	0	
3	2	US-2		-117.5334722	47.643175	1	53	63	9					27	2	0	
1	2	1ST AVE	COLUMBIA SR	-122.334175	47.60174444	3	53	12	12	8				24	1	0	
1	2	1ST AVE	COLUMBIA SR	-122.334175	47.60174444	1	53	12						49	2	0	
1	2	1ST AVE	COLUMBIA SR	-122.334175	47.60174444	2	53	12						37	1	0	
1	2	I-5		-122.9751333	46.70995833	1	53	8						19	2	0	
3	1	I-90		-121.0397611	47.18249167	2	53	12	63	25	1	12		21	1	0	
3	1	I-90		-121.0397611	47.18249167	1	53	12	12	63	25	64	48	24	1	0	
3	1	I-90		-121.0397611	47.18249167	2	53	12	12					42	2	0	
4	2	SR-26		-118.576325	46.79465	2	53	68	12					20	2	7	
4	2	SR-26		-118.576325	46.79465	1	53	12						60	1	0	
2	2	SR-3		-122.9179833	47.31120556	1	51	68	12					23	1	0	
2	2	SR-3		-122.9179833	47.31120556	1	53	12						18	2	7	
4	2	CR-S BROOKS RD		-117.6902167	47.625875	2	53	10						45	2	1	
1	1	I-5		-122.3091917	47.54683611	2	53	12						46	2	7	
1	1	I-5		-122.3091917	47.54683611	3	53	12						29	1	0	
4	2	CR-RESERVATION RD		-122.6971917	47.02611111	1	53	68	68	63	38	1		32	1	0	
1	2	S MARKET BLV	4TH ST	-122.96055	46.65916389	2	53	8						22	2	0	
1	2	CR-CANYON RI	104TH ST E	-122.3571611	47.16114167	1	53	68	68	1				24	1	7	
1	2	I-90		-120.6309167	47.04091111	2	53	12	63	24				21	2	7	
1	2	I-90		-120.6309167	47.04091111	3	18	12						49	1	0	
1	2	I-5		-122.5424556	47.127125	2	53	8						25	1	0	
1	2	SR-543	BOBLETT ST	-122.7349472	48.99055	9	95	12						32	1	0	
1	2	SR-543	BOBLETT ST	-122.7349472	48.99055	9	95	12	12					40	1	7	
1	2	SR-543	BOBLETT ST	-122.7349472	48.99055	3	53	12						34	1	0	
1	1	I-405		-122.2493472	47.46273056	1	53	12	64	25				35	1	7	
1	1	I-405		-122.2493472	47.46273056	1	53	12						38	1	0	
2	2	176TH ST SW		-122.2937917	47.83918333	1	53	8						53	2	0	
2	2	SR-16		-122.6245306	47.46971389	1	53	14						29	1	0	
3	1	US-12	W POWERHOU	-120.5942306	46.63881389	1	53	12						25	1	7	
3	1	US-12	W POWERHOU	-120.5942306	46.63881389	3	53	12	68	1	12			33	1	0	
3	1	US-12	W POWERHOU	-120.5942306	46.63881389	1	53	12						58	2	0	

Reverse Geocoding

Thought-Process

```
# Running the Nominatin class using a timeout function in order to prev
locator = Nominatim(user_agent="geocoder", timeout = 10)

# Read the CSV data.
df = pd.read_csv(csv_file)

# Initializing a few lists for later usage.
zip_data = []
zip_header = ['zip']

# Iterating through each row within the dataframe,
# Reverse Geocode the coordinates and write into TXT file.
for a, b in df.iterrows():

    coordinates = (b['y'], b['x'])
    addr = locator.reverse(coordinates)

    # If there is a postcode and an address, append the zipcode.
    if ('address' in addr.raw and 'postcode' in addr.raw['address']):
        zipcode = (addr.raw['address']['postcode'])
        zip_data.append(zipcode)
    # Otherwise, append 0 as the zipcode.
    else:
        zip_data.append('0')

# Create a CSV file with writing permissions.
fp = open('reverse_geo.csv', 'w', newline='')

# Create a CSV writer for 'fp'.
writer = csv.writer(fp)

# Create one header row called 'zip'.
writer.writerow(zip_header)

for x in zip_data:
    # Write each row into CSV.
    writer.writerow([x])

# Close the file.
fp.close()
```



Thought-Process

Analyze Data

- Using variety of tools to help us perform most of our analysis on dataset.

Use Spread Sheet Commands

- Compare zip codes to crash zip codes to categorize crashes within/outside community.

Provide Visualization Through Tableau:

- Zip Codes that produced the most dangerous drivers
- Sequence of events were most common for crashes (within/outside community)

Data Analysis

Thought-Process

Solution

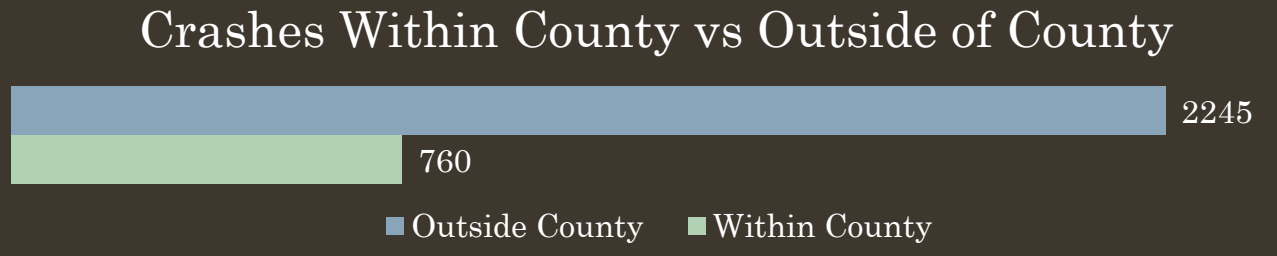
Solution

1. Among drivers involved in fatal crashes, what proportion are involved in crashes in communities where they live?

a. Are there differences in the types of crashes and behavior factors in those crashes among "residents" versus those deemed to be not "from" the area?

What We Learned:

- About 1/3 of crashes happen 'within the same zip code.'
- Main sequence of events are the same for both 'within county' and 'outside county'.
- Crashes were more likely to occur at night or in dark conditions out of community.



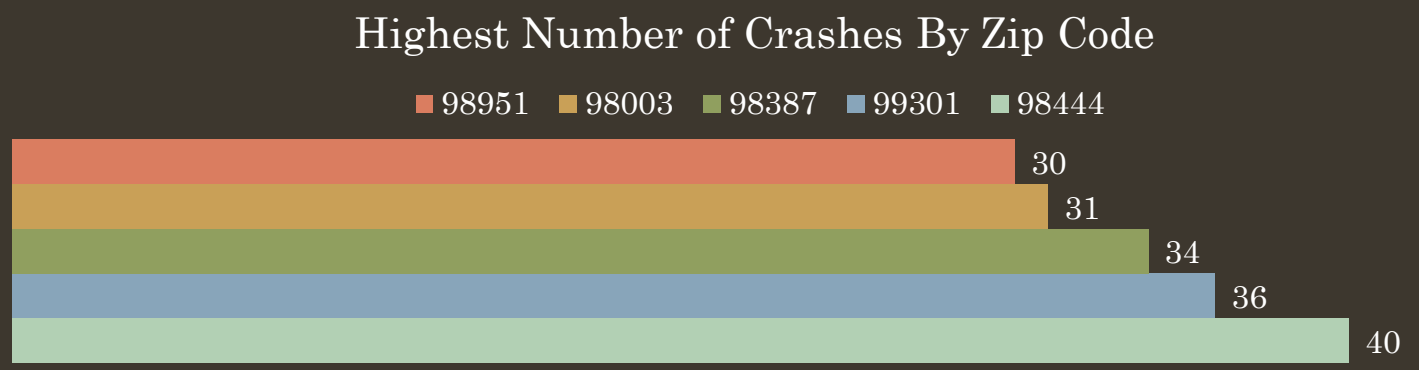
Seq. of Events	Within County	Outside County
Motor Vehicle and Transport	338	938
Run Off Road - Right	153	444
Cross Centerline	140	442
Collision with Pedestrian	89	391

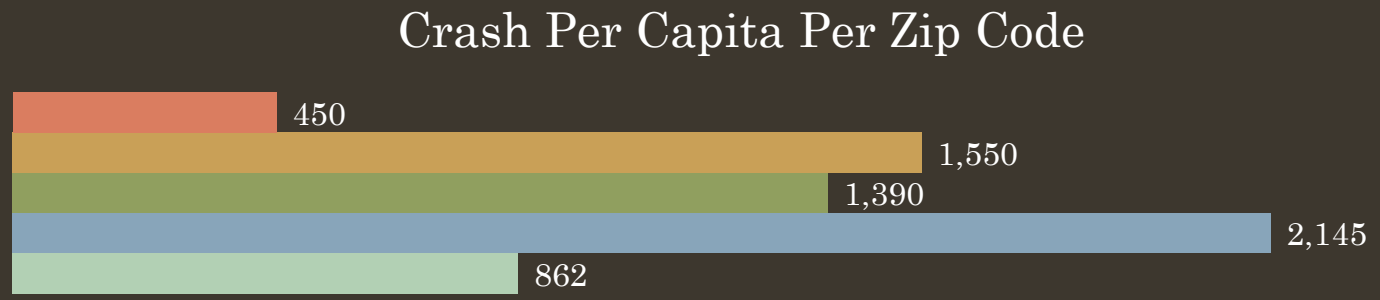
Solutions Cont.

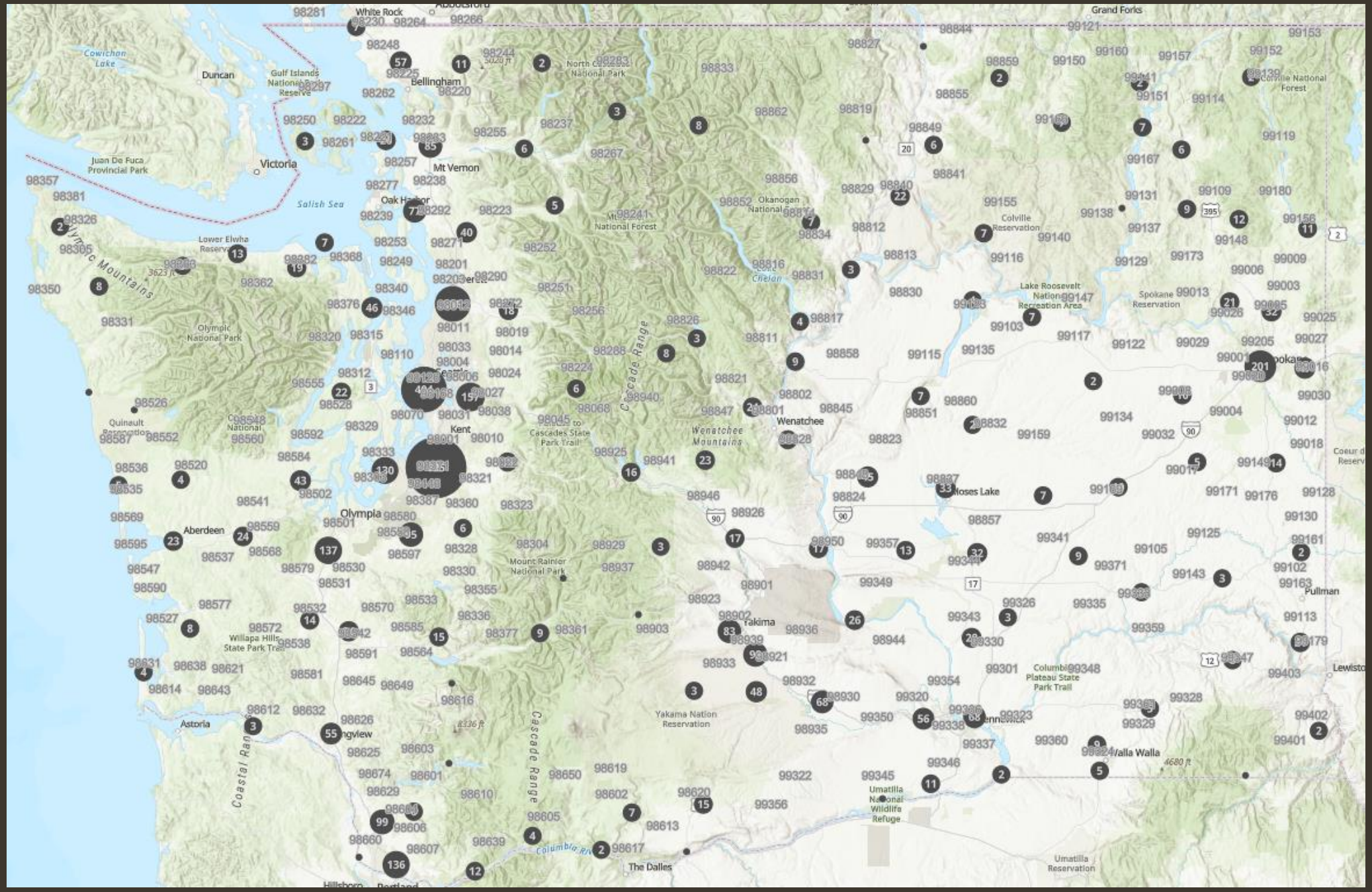
2. Are there specific resident ZIP Codes that tend to produce higher-risk drivers that are involved in fatal crashes at a higher rate?

a. What are the population demographics of these high-risk driver producing ZIP Codes?

- Tacoma
- Pasco
- Spanaway
- Federal Way
- Wapato





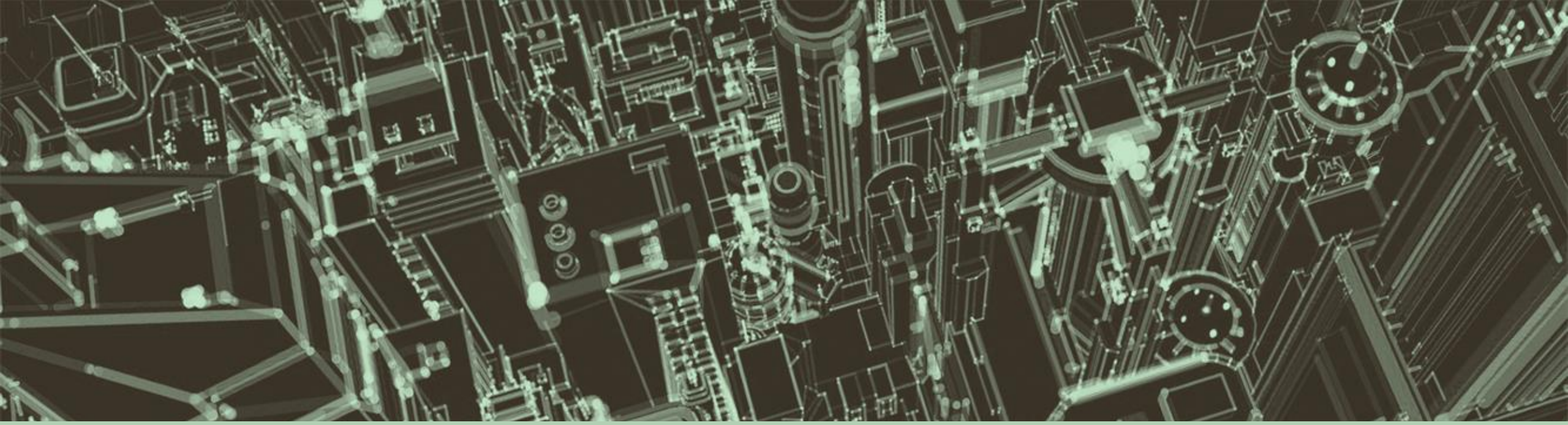


Washington Fatal Crashes Data Files – Data Analytics Level 4

Future Considerations

Create A Prediction Model

Analyze Pre-COVID vs Post-COVID



Thank you!