IC23: INFORMATION & ACCESS FOR ALL

The Washington Traffic Safety Commission Dataset

Making sense of senseless road deaths.

Team 23050:
Nicole Schneider
Kent O'Sullivan
Harsh Patel

04 Mar 2023

Key Questions:

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

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

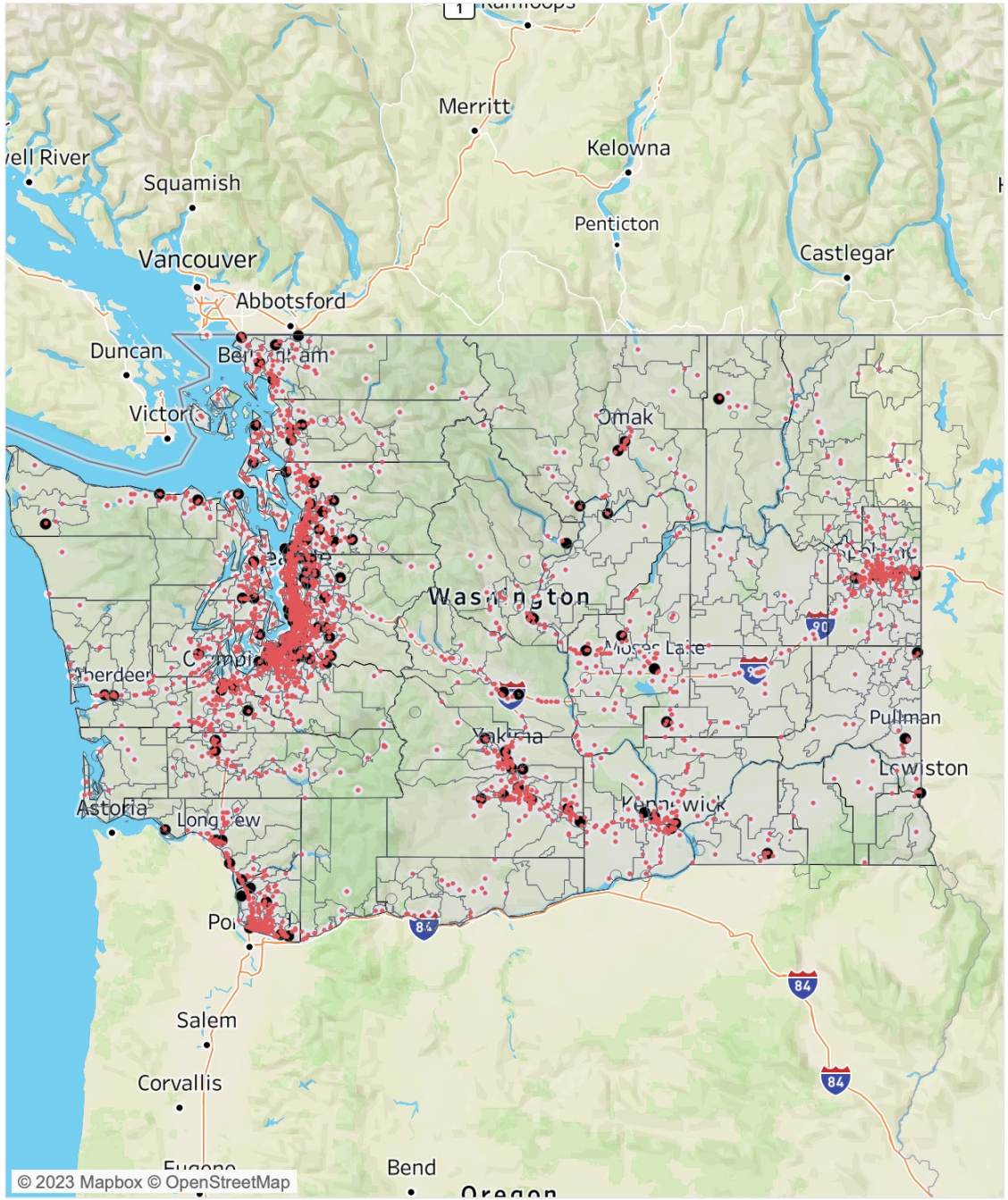
3. Is there any evidence that the way Washington drivers adapted during the COVID-19 pandemic has impacted road fatalities

Key Requirements:

1. The work and results should be interpretable by people who are domain experts, not ”Programming People”

2. The work and results should be interpretable by people who are domain experts, not ”Data People”

PROBLEM OVERVIEW



Washington State:

- [ ] Area: ~ 71,000 Sq mi
- [ ] Terrain: Mountainous
- [ ] Population: ~ 7.9 Million
  - [ ] 39 Counties
  - [ ] 281 Cities and towns
  - [ ] 592 ZIP Codes

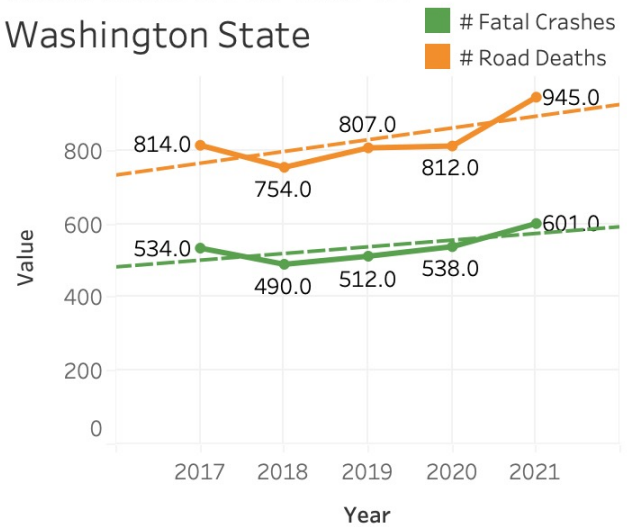
The Data:

- [ ] ~ 4000 Records
- [ ] Temporal bound: *2017-2021*
- [ ] Spatial bound: *Washington State**
- [ ] Data: *160 fields (encoded), ZIP missing.*

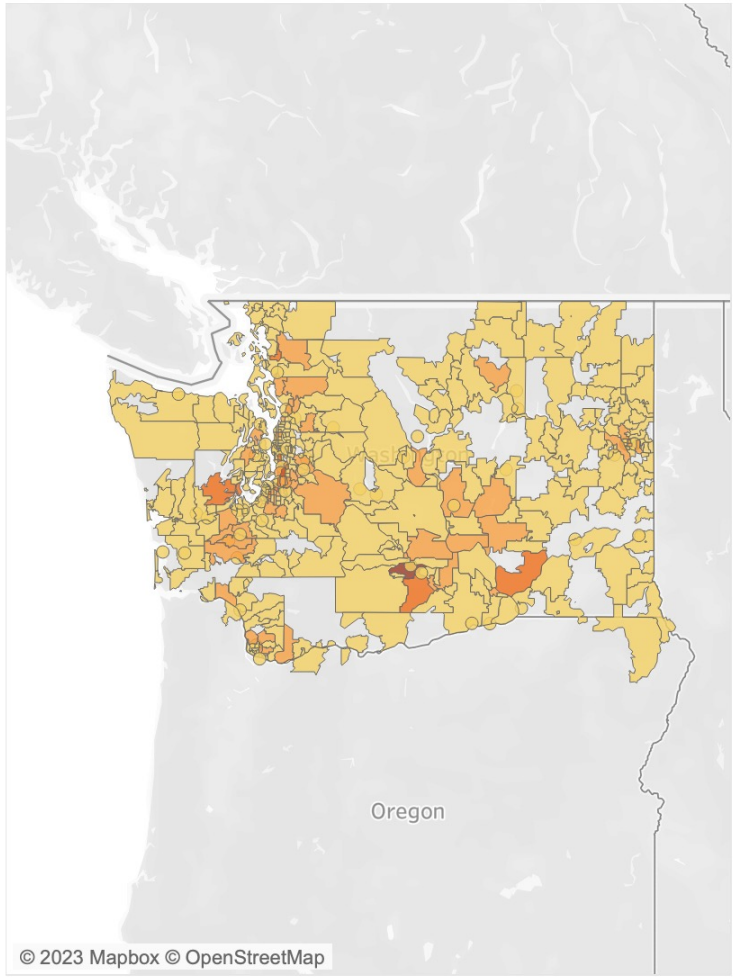
The Problem:

- [ ] Which crashes are in which ZIP codes?

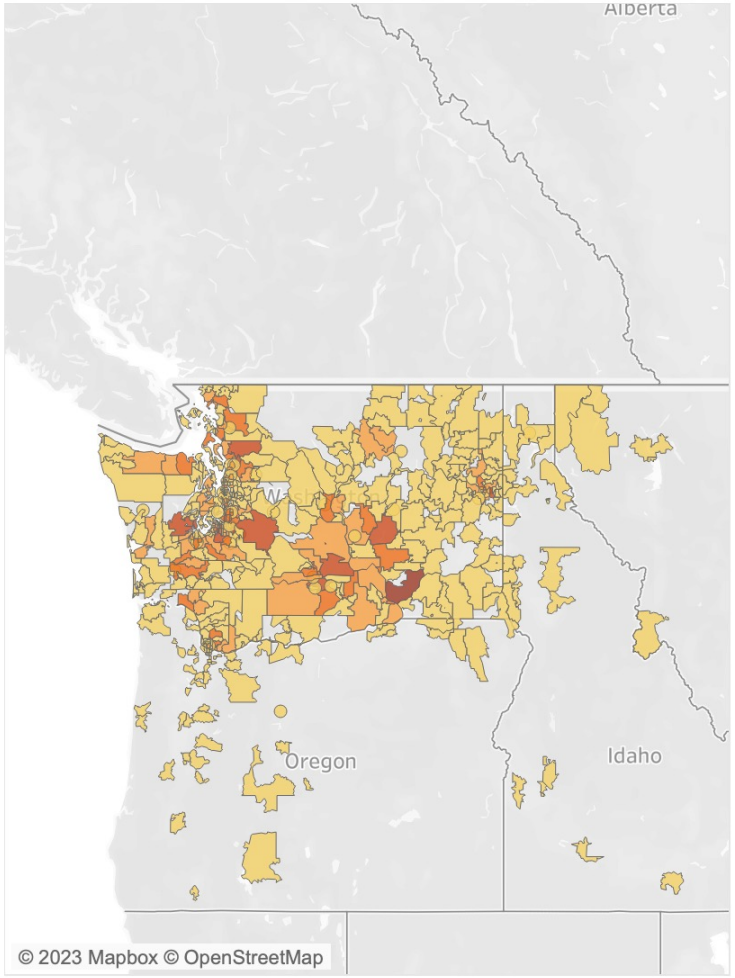
Road Deaths Per Year in Washington State



Heatmap of Crash ZIPs



Heatmap of Driver Origins



% of Total Count of Case

0.024% 1.646%

% of Total Count of Case

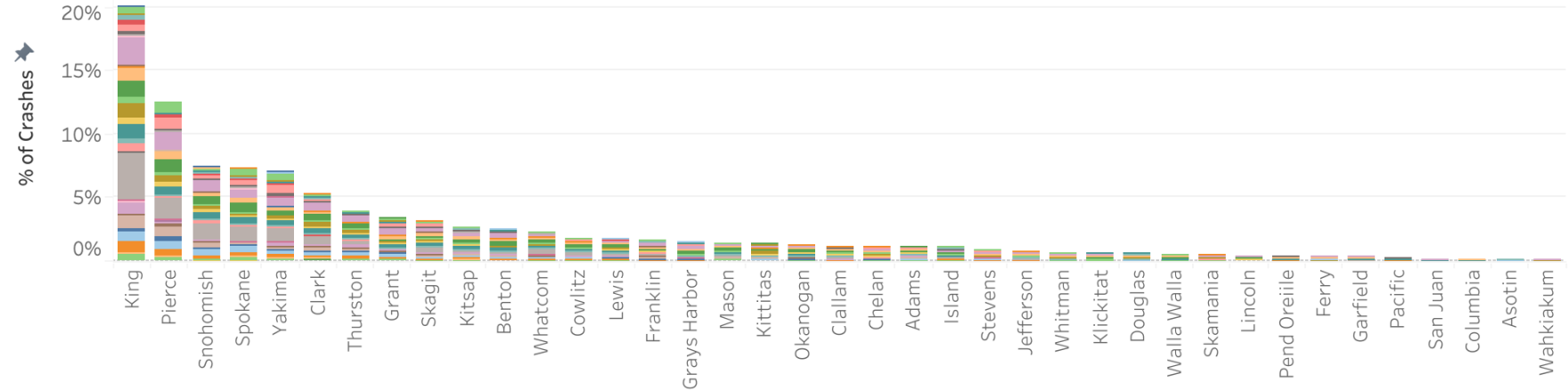
0.024% 1.137%

KEY TAKEAWAYS:

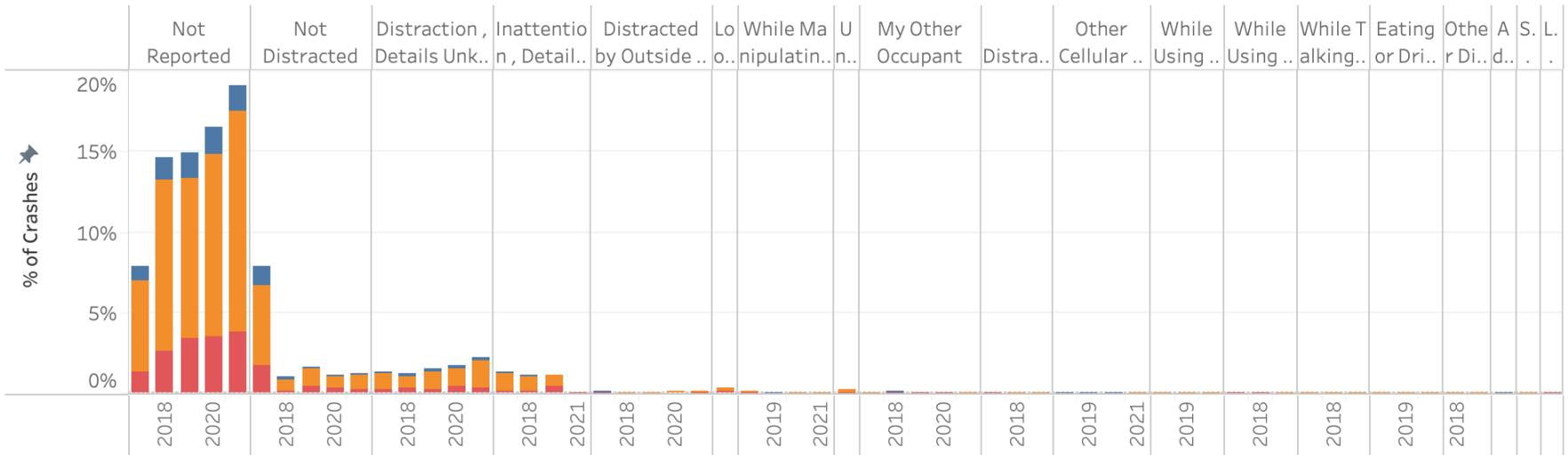
- Most crashes happen close to population centers.

- Drivers from out-of-state are a negligible number of crashes.

Critical Events by County



Distraction Factors by Year



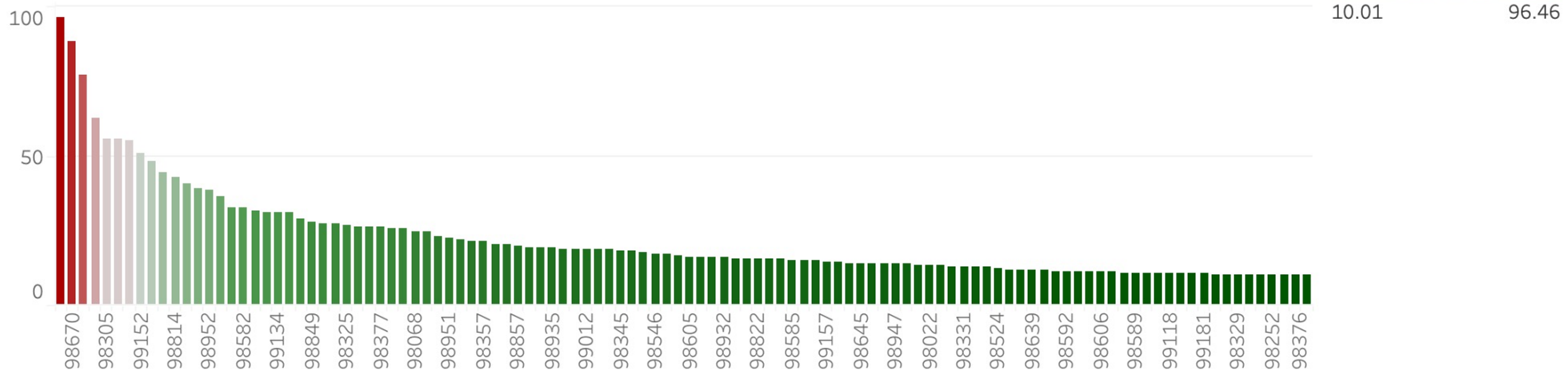
KEY TAKEAWAYS:

Most crashes are concentrated around population centers.

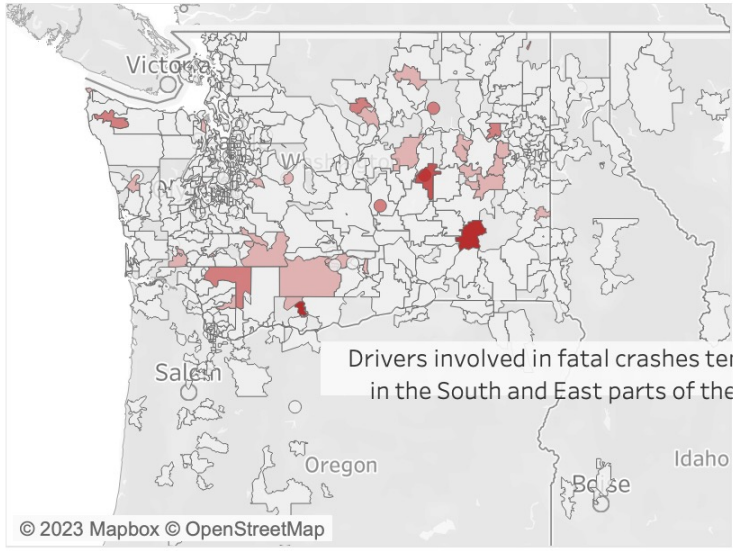
Most Counties have comporable causes of crashes.

We don’t have enough data about causes of distractions (reverse-survivorship bias)

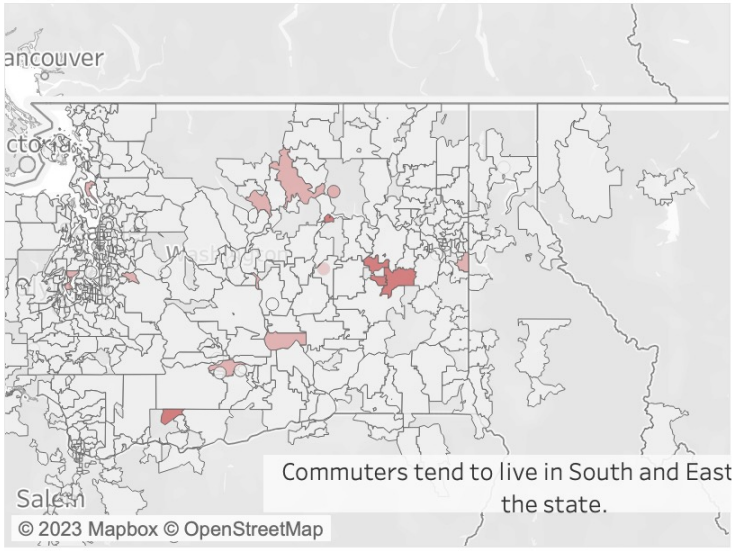
ZIPs Producing Bad Drivers



Heatmap of ZIPs Producing Bad Drivers



Heatmap of ZIPs Producing Commuters

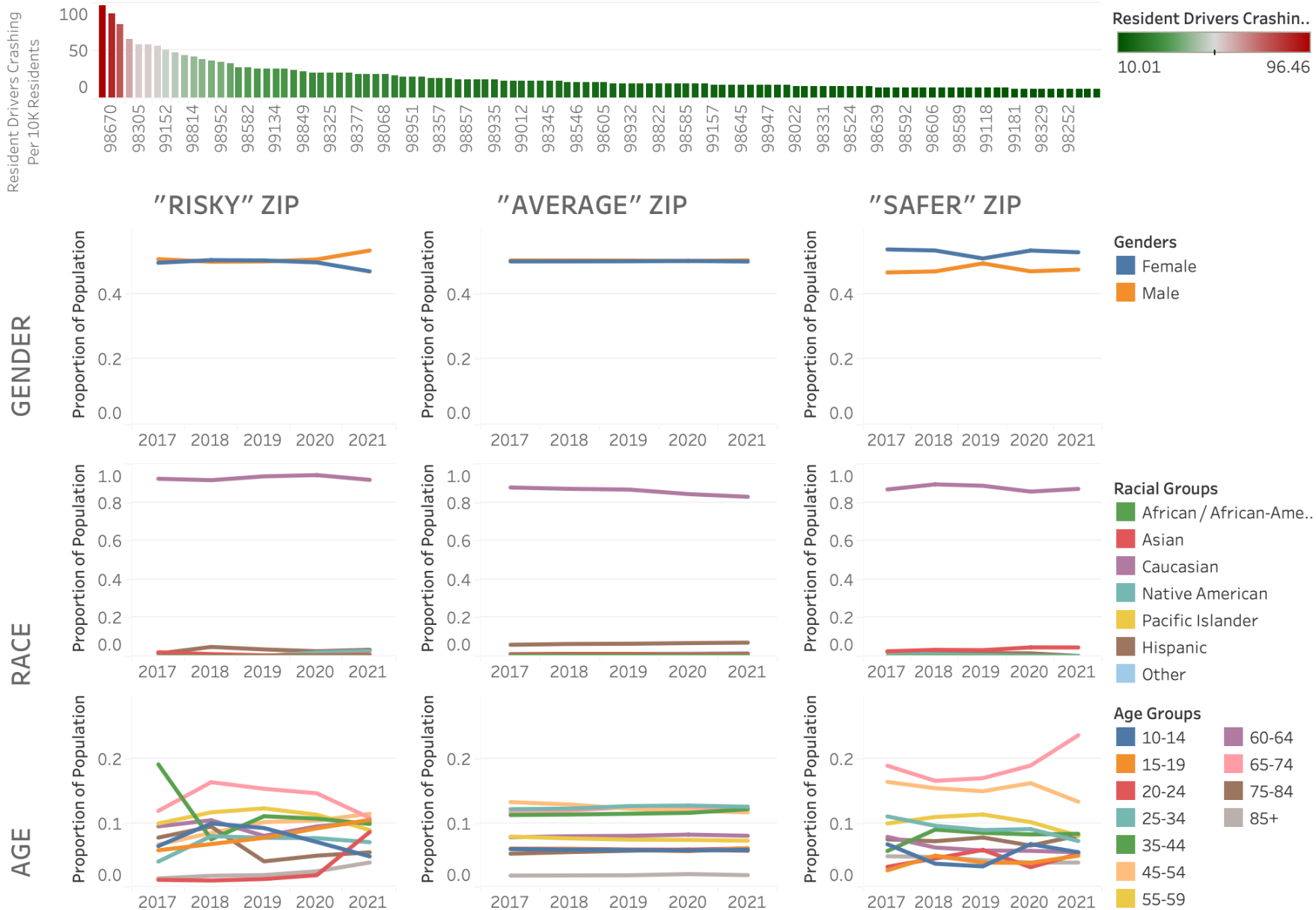


KEY TAKEAWAYS:

Population density and ZIP code size distort the origins of 'bad' drivers.

'Riskier' drivers tend to live in the south and east of the state.

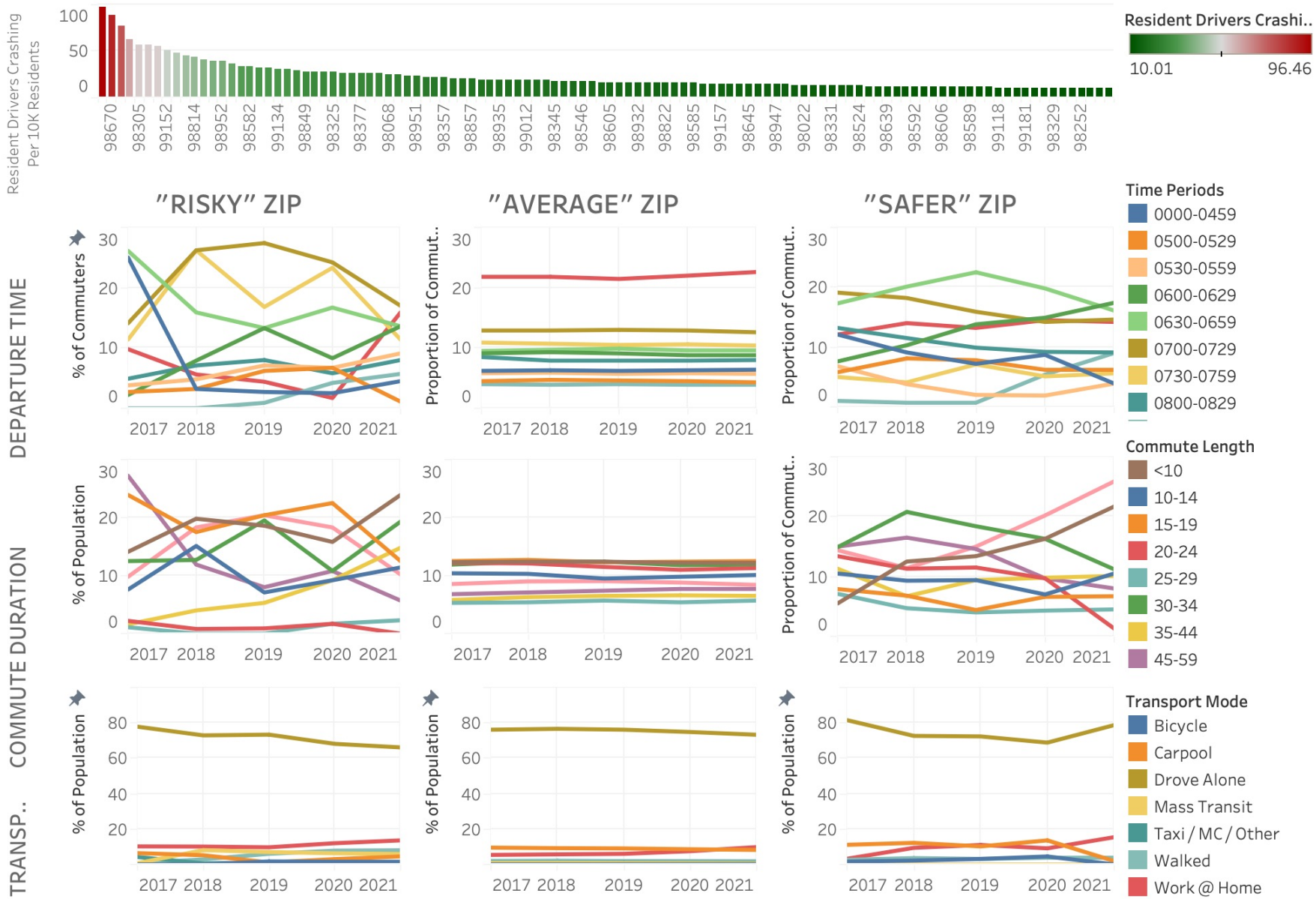
'Commuters' tend to live in the south and east of the state. ..



KEY TAKEAWAYS:

No real evidence for demographic causality on these dimensions

Younger Male populations have some correlation with risky behaviour



KEY TAKEAWAYS

No real causal commuting factors

Non-standard hours, short commutes and use of public transport in ZIP may indicate wealth-related factors

SUMMARY

Our Probem

Why are road fatalities increasing, can we identify riskier communities and is it related to COVID-19?

Our Process

Make interpretable workbook so that we can 'see' the data in space and time.
Enrich the data with geocoding and census data about demographics and commuting.

Our Findings

There's no real evidence that this is anything more than random variance.
The locations and causes of crashes are fairly uniformly distributed, with concentration around highly populated areas.
We identified no strong demographic factors.
The causes and locations are (mostly) consistent over time .

Our Next Steps

Incorportate non-fatal crash data.
Incrorporate economic status data
Hand the tools over to the experts