Autonomous Truck Deployment – Impact of Road Construction

July 2025

Presented by: Salah Uddin Momtaz

Presented to: Aurora

Journey Through the Data

Setting the Stage: AV Trucks Meets Reality

Uncovering the Hidden Patterns

Emerging City Construction Landscape

Cities to Routes: Tires on Pavement

Predicting Construction Chaos

Analysis to Action

Discussion: Path Forward

Setting the Stage: Where AV Trucks Meet Reality

Promise vs. Daily Roadblocks

$168B

Annual Savings promised by autonomous trucks

8.8K

Construction causing daily disruptions in 2021

Autonomous fleet needs to navigate this maze of construction – minimizing costly delays, route failures, customer dissatisfaction

Data Sources Behind the Story

US Road Construction and Closures (2016 - 2021)

Route Line

Census Data

Road Network GIS

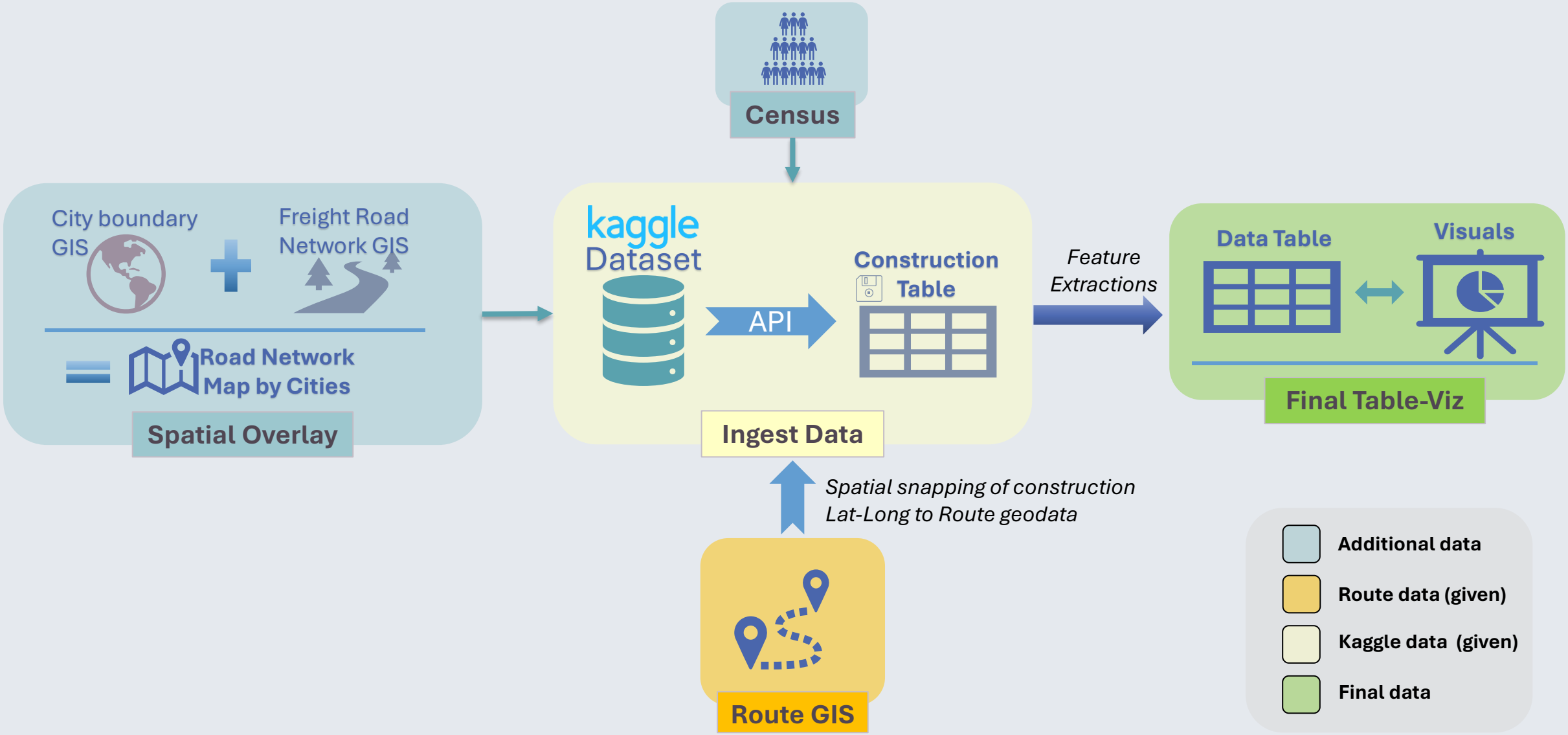
Uncovering the Hidden Patterns

What 6.2 Million Construction Projects Revealed

Summary

- Data: US road construction data (2016-2021)
- Critical insights for AV deployment strategy
- Identifying high-impact zones and routes
- Predict the impact of constructions

Connecting Construction - Route

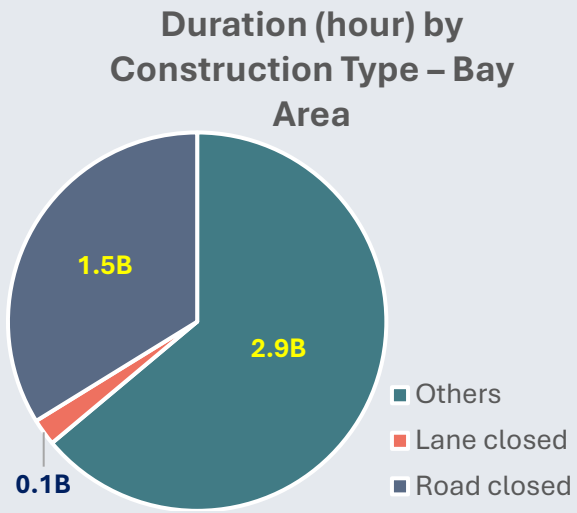


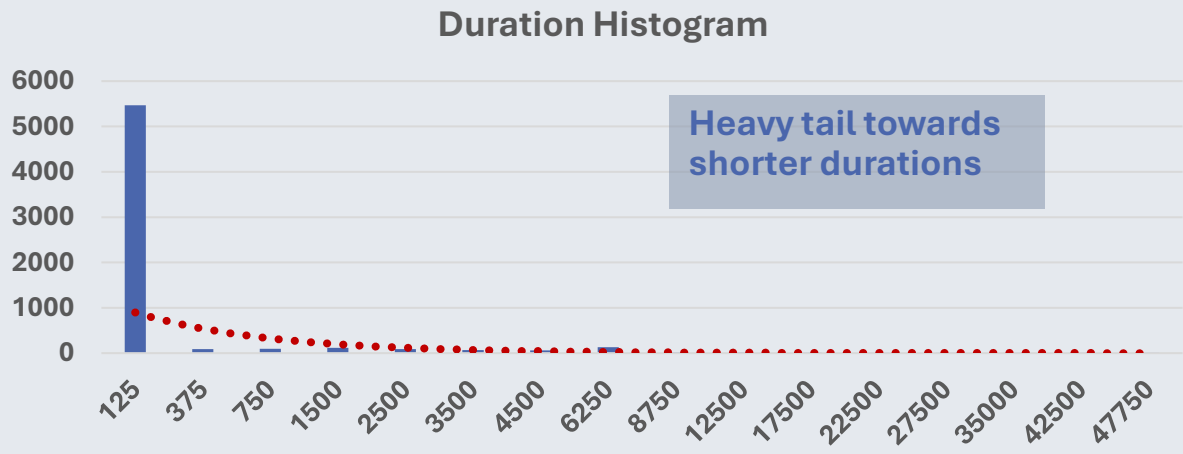
First Discovery: Not All Disruptions Are Equal

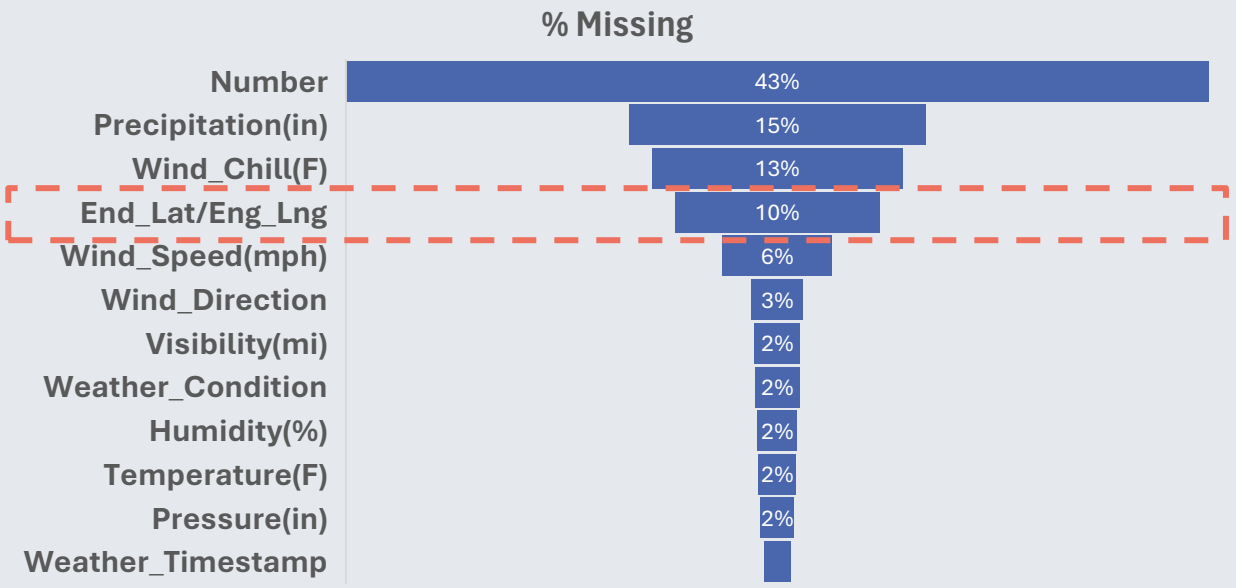
Summary Steps

- Feature Extractions
  - Duration → From start and end time
  - Road/Lane Closure → From “Description”

- Distribution & Missing Values







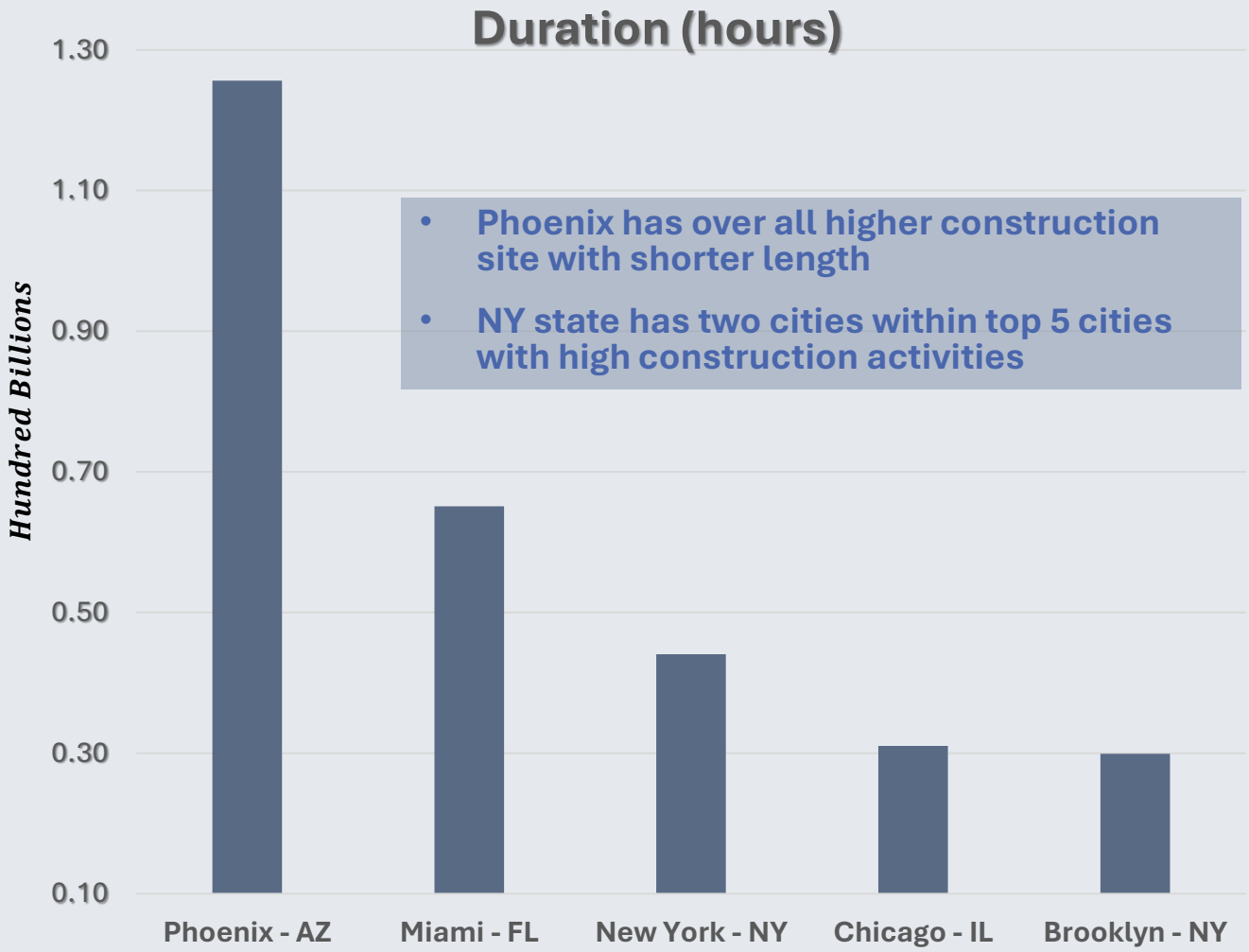
The City Construction Landscape Emerges

Phoenix Paradox: Most Construction, Shortest Delays

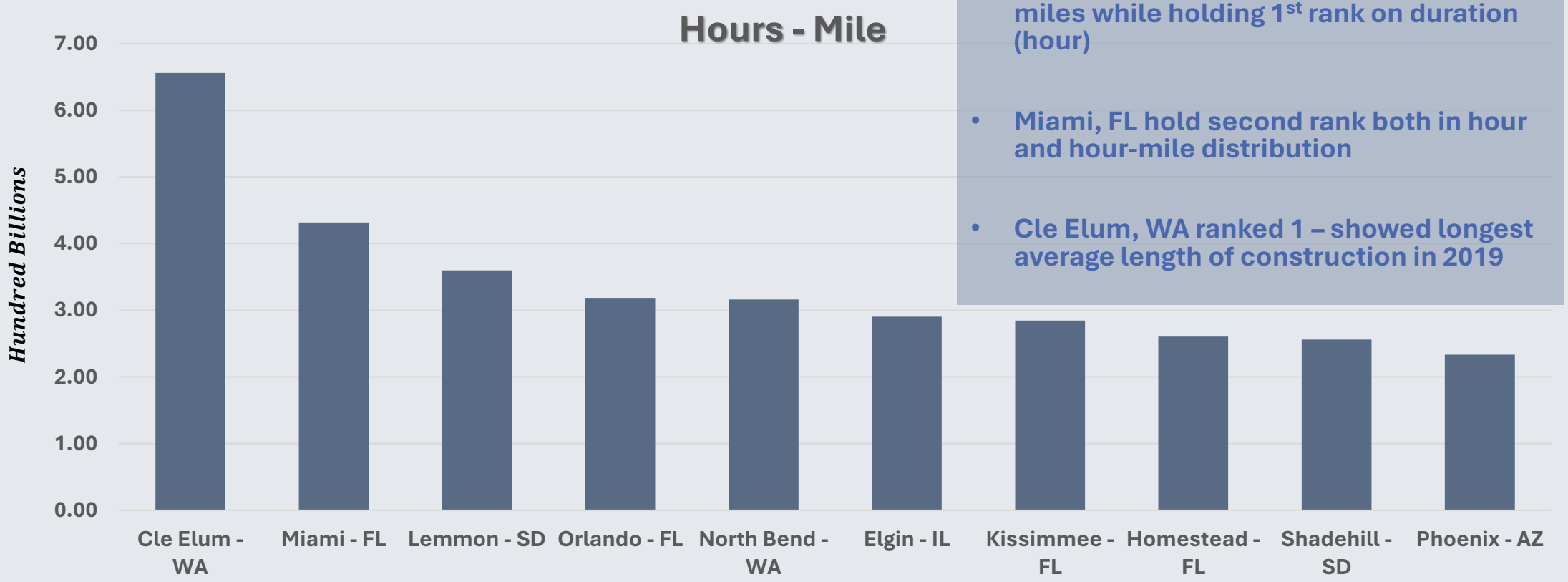
Top Cities Over 6 years

Year	🏆 Top City	Million Hours	Sites	Avg. Distance (miles)
2016	Middletown, DE	10.4	299	0.9
2017	York, PA	10.3	841	1.8
2018	Winnfield, LA	9.8	461	1.9
2019	Cle Elum, WA	**17.8**	1,393	**4.0**
2020	**Phoenix, AZ**	11.9	34,240	0.2
2021	**Phoenix, AZ**	10.3	**212,356**	0.2

- Phoenix has higher number of construction sites with shorter length
- Cle Elum has long stretched sites

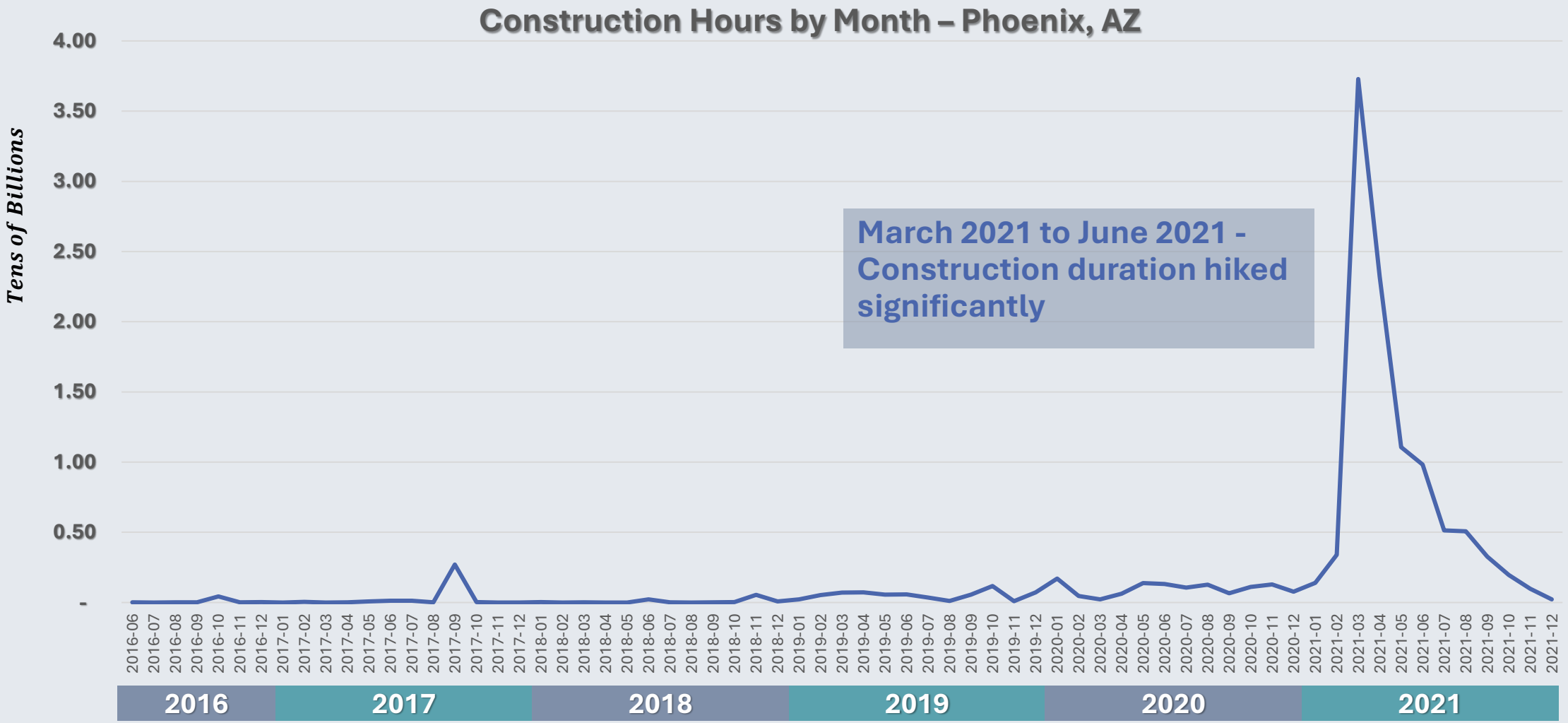


More Phoenix Paradox: Ranks #1 in Hours but #10 in Impact

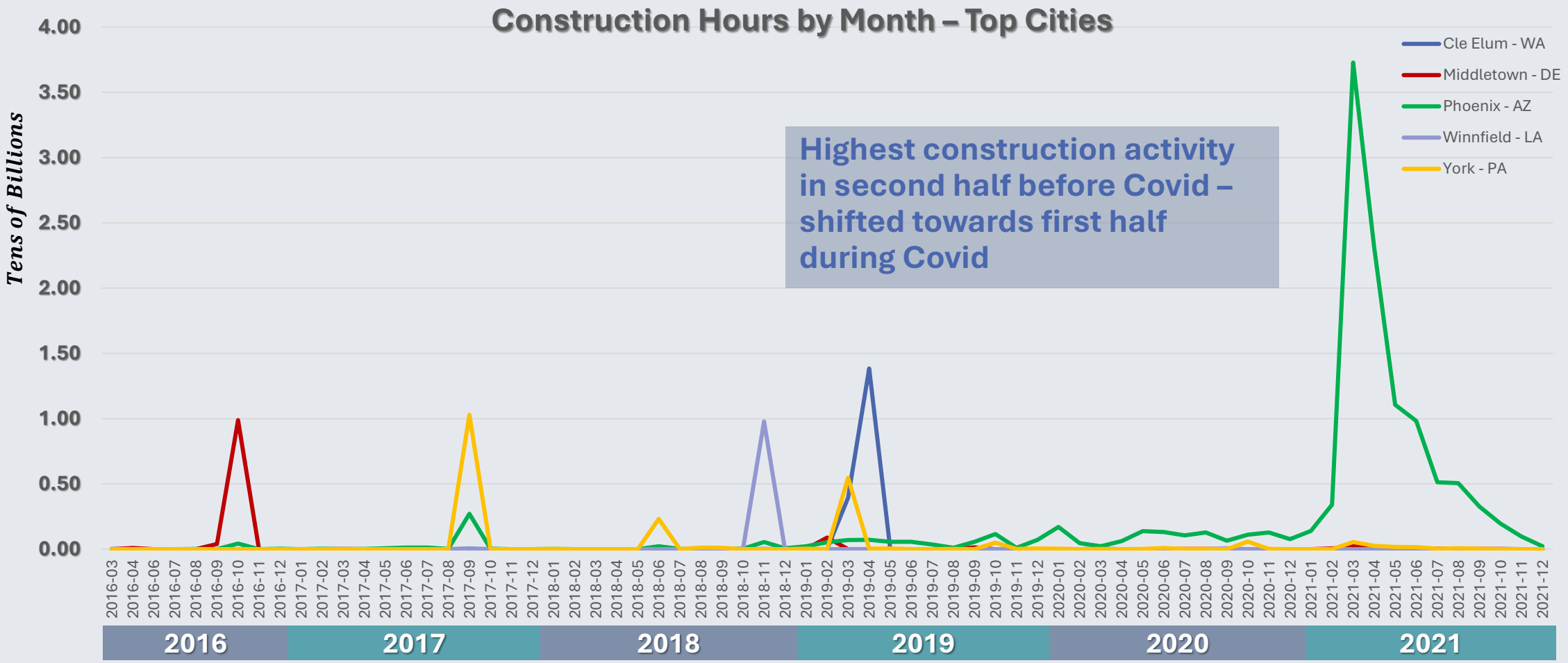


- Phoenix, AZ ranked 10th in terms of Hour-miles while holding 1st rank on duration (hour)
- Miami, FL hold second rank both in hour and hour-mile distribution
- Cle Elum, WA ranked 1 – showed longest average length of construction in 2019

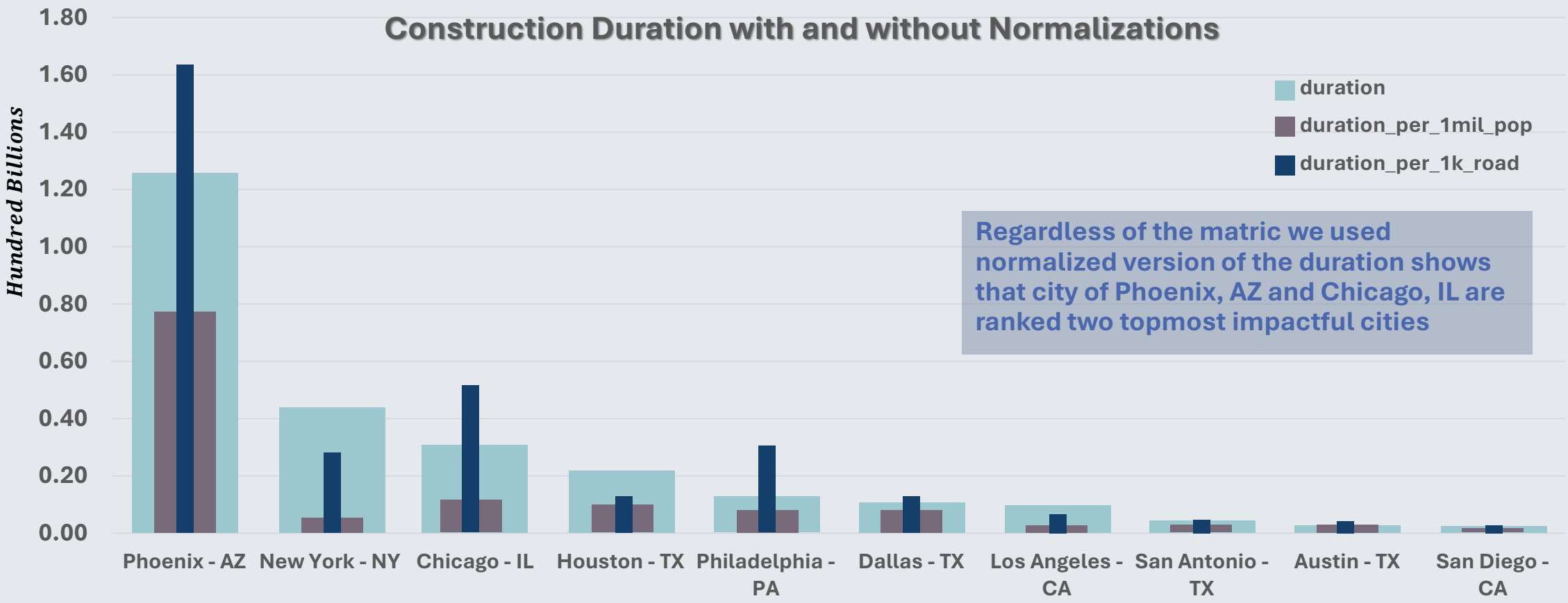
March 2021: Construction Patterns Change



COVID's Gift: Shifting Construction Seasons

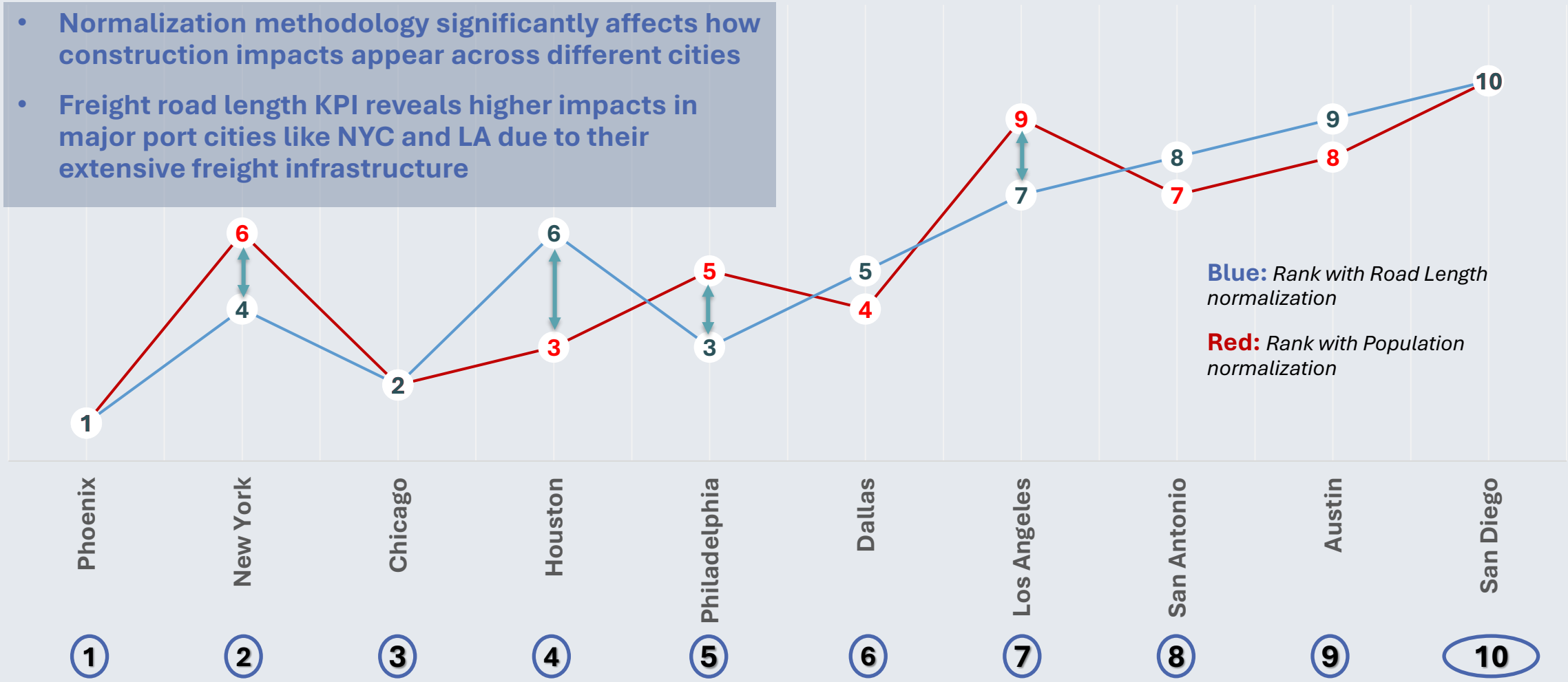


Phoenix & Chicago: The Construction Giants



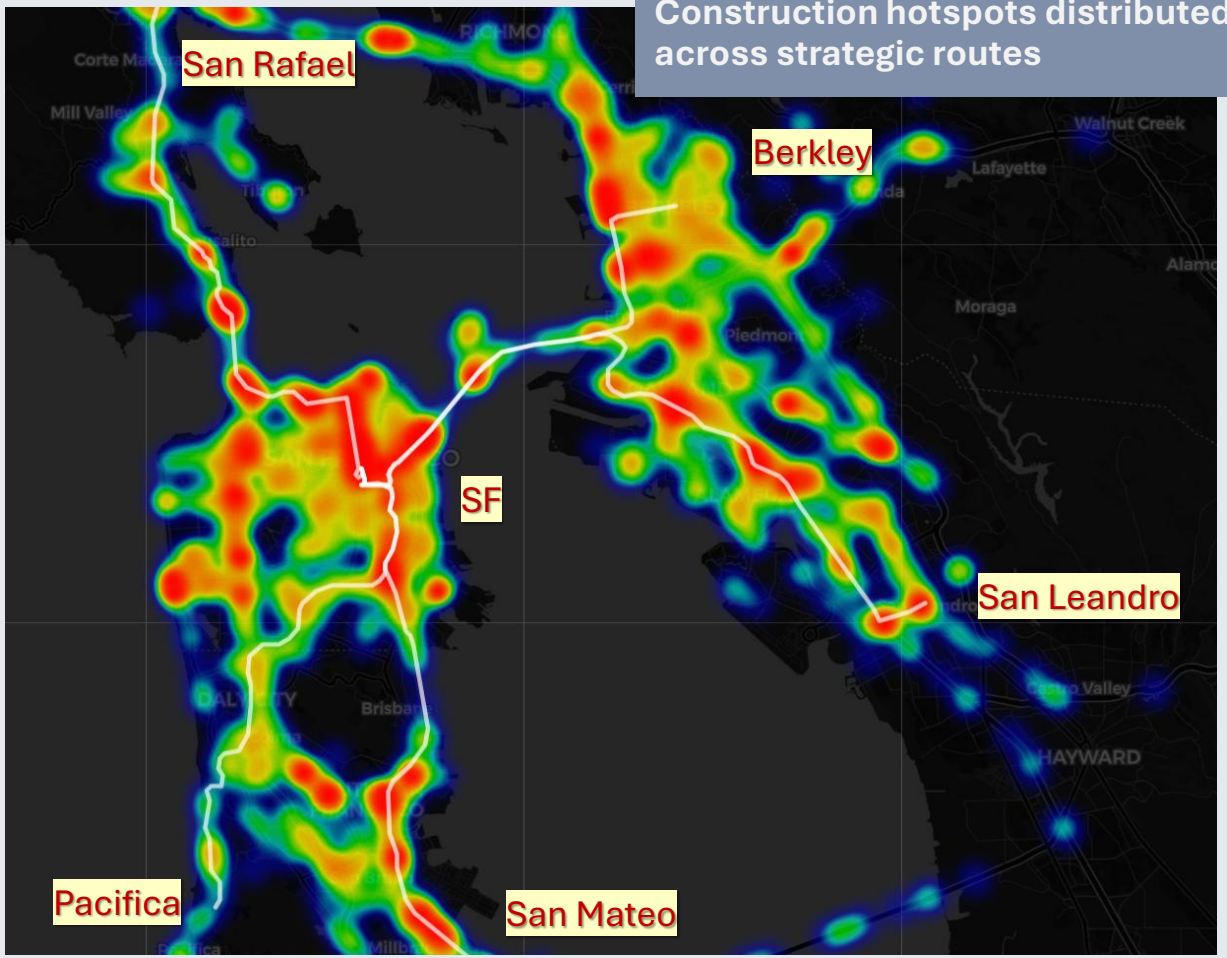
Normalization Method Changes Everything

- Normalization methodology significantly affects how construction impacts appear across different cities
- Freight road length KPI reveals higher impacts in major port cities like NYC and LA due to their extensive freight infrastructure



From Cities to Routes: Where Rubber Meets Road

Heat Map Reveals: One Route Stands Apart



Geospatial Analysis Summary

📊 32K+ total construction events in Bay Area

📏 20m radius threshold for spatial snapping to matching construction sites with Routes

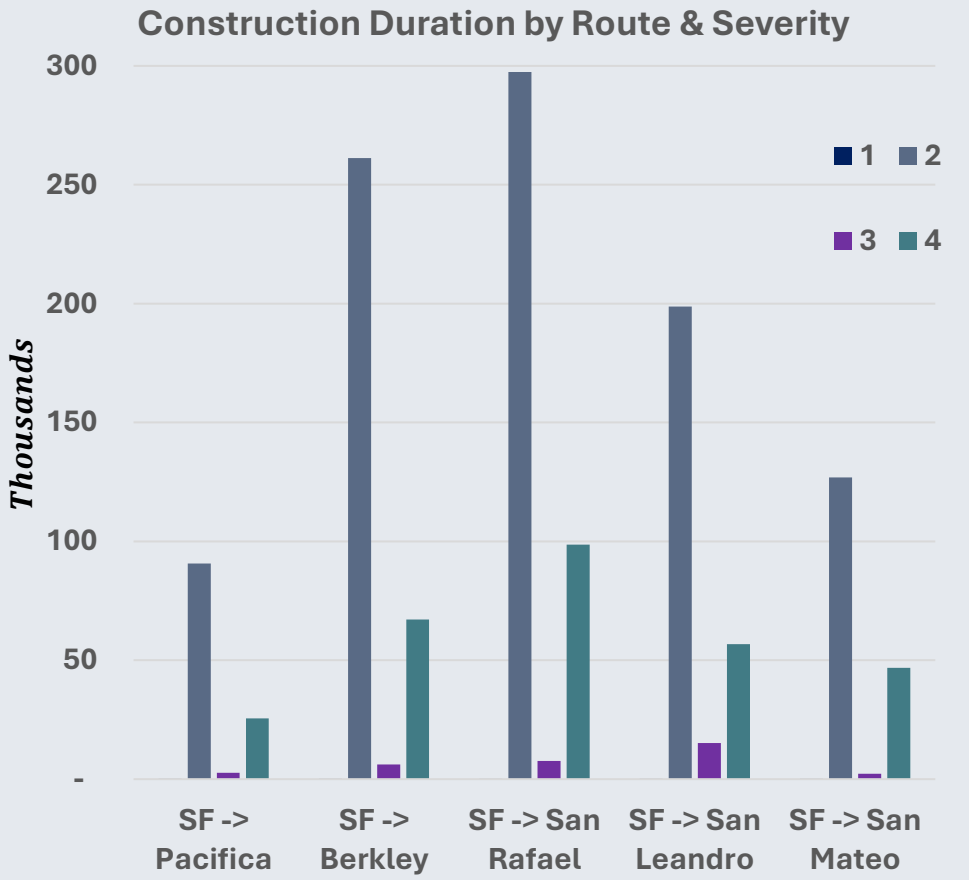
🎯 9K+ events directly impacting routes

SF → Pacifica shows visibly lower construction density

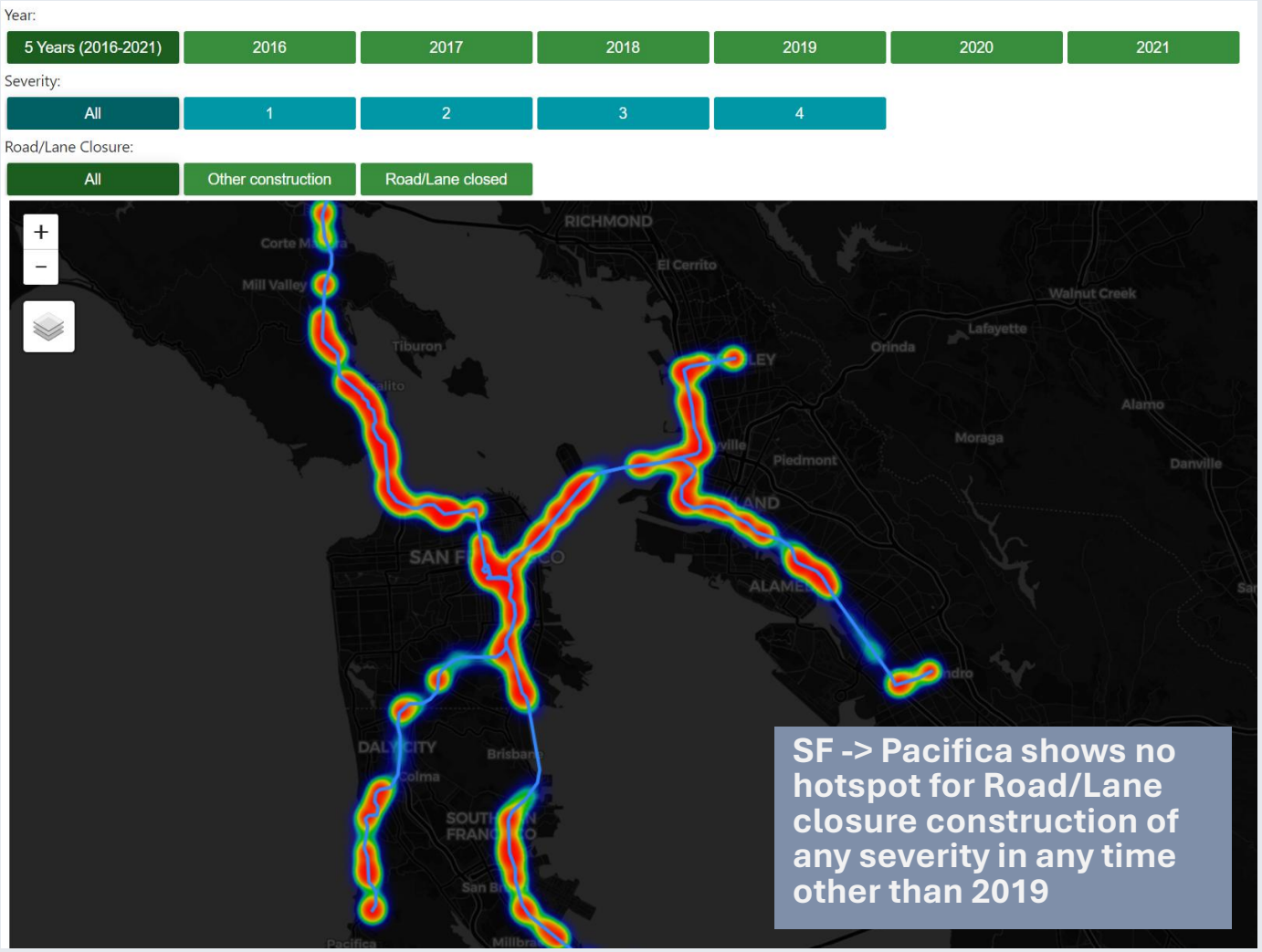
SF → San Rafael traverses high-intensity zones

Construction hotspots **concentrate around urban core**

SF-Pacifica: The Clear Winner for AV Testing



SF -> Pacifica shows lowest number of construction with higher severity



SF -> Pacifica shows no hotspot for Road/Lane closure construction of any severity in any time other than 2019

Quantifying Risk: From Best (1) to Worst (0)

Risk Scoring Methodology

Feature Weights (by Standard Deviation)

- 🚧 Construction Duration: **96.5%**
- 📏 Distance Impact: **1.3%**
- ⚠️ Severity Weight: **1.4%**
- 🚦 Road Closure Events: **0.8%**

Risk Score – Developed by Inverse Normalization

Route	Risk Score	Rank	Hours	Road Closure	Distance (mi)
SF → Pacifica	**1.00**	**1**	**119K**	**1.1K**	**0.9K**
SF → San Mateo	0.80	2	176K	1.6K	1.2K
SF → San Leandro	0.45	3	271K	3.1K	4.2K
SF → Berkeley	0.24	4	335K	3.2K	3.9K
SF → San Rafael	**0.01**	**5**	**404K**	**2.9K**	**2.7K**

Can We Predict Construction Chaos?

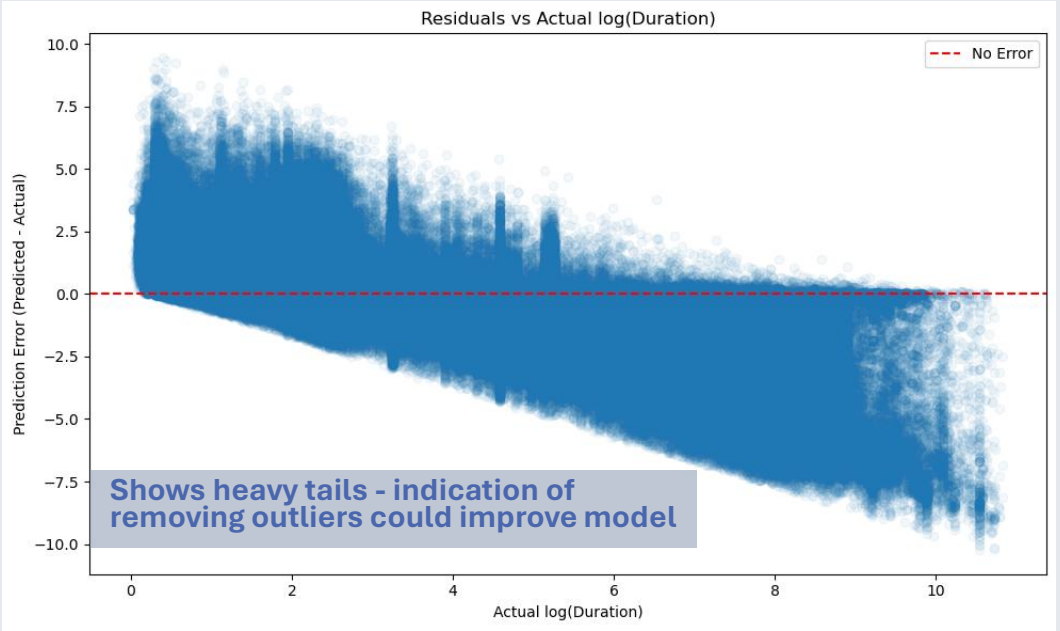
Random Forest Wins

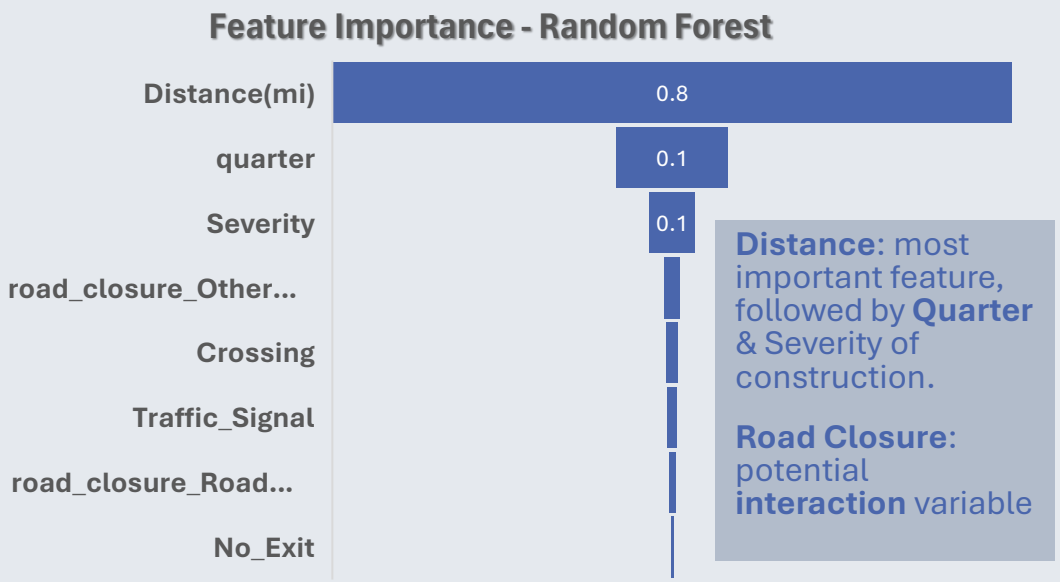
Log Linear Regression Model

- Training, Test data split 80% - 20%
- Performs poorly
  - MAE: 1.645
  - RMSE: 2.132
  - R^2: 0.076

Random Forest Regressor

- Training, Test data split 80% - 20%
- Good enough for strategic/sketch planning, needs improvement for daily operations
  - MAE: 1.2344
  - RMSE: 1.851
  - R^2: 0.303





The Path to Improve Predictability

Current Performance Benchmark: 30% R^2 - insights for strategic decisions

Target Enhancement: 50%+ R^2 - could aid operational automations and dynamic deployments

Model Enhancement Roadmap

Outlier Treatment

- *Remove top or bottom 5%*
- *Down-weight the outliers*

Interaction Terms

- *Road closure, Severity, Quarters and distance interactions could work*

Ensemble Methods

- *Evaluate Geospatial Models*
- *Segmented model for different spatial levels*

Models Helps to Quantify Actions

Predict Duration → Estimate Delay → Estimate Operating Cost → Score by Cost → Zore-cone – Minimal cost Route

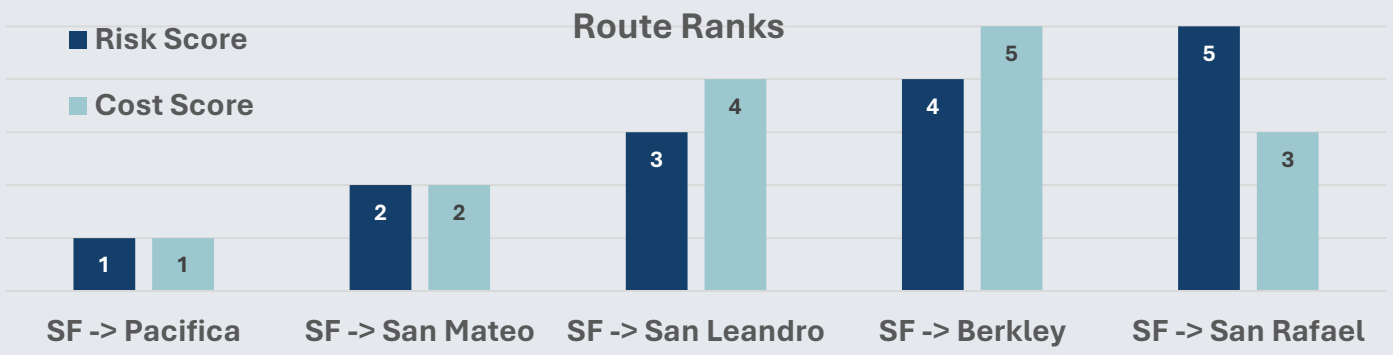
Route	Avg Duration (Predicted)	Annual Events	Delay Factor	Exposure Factor	Est. Delay (hours)	Operating Cost	Rank by Cost
SF → Pacifica	**43.7**	**240**	**0.28**	**0.28**	**820**	**$125K**	**1**
SF → San Mateo	38.2	341	0.29	0.43	1,620	$245K	2
SF → San Leandro	47.0	603	0.22	0.44	2,640	$410K	4
SF → Berkeley	**43.8**	**652**	**0.20**	**0.50**	**2,860**	**$430K**	**5**
SF → San Rafael	44.7	618	0.22	0.34	2,070	$310K	3

Saves $300K per truck per year

Delay Factor →
Construction type Factor * Proportion
Construction type

- **Road Close: 1**
- **Lane Close: 0.3**
- **Others: 0.1**

Exposure Factor →
Hour – Mile Normalization in [0 - 0.5]



Turning Analysis into Action

Findings & Actions: Roadmap Forward

Key Findings

- **Phoenix, AZ** emerges as highest risk city with 212K+ construction sites but shorter duration projects
- **COVID-19 (/*Election*) caused a shift in construction patterns from Q3-Q4 to Q1-Q2**, requiring adaptive route planning
- **SF → Pacifica** route shows **minimal construction risk**, ideal for initial deployment
- **Random Forest model** achieves over **30% accuracy** with **distance** as the primary predictor

Recommendations

- Implement **dynamic routing algorithms** for high-impact cities like **Phoenix** and **Chicago**
- Develop **seasonal deployment strategies** accounting for **Q1 construction peaks**
- **Prioritize low-risk & low-cost routes** like SF-Pacifica for initial autonomous deployment
- Enhance **model accuracy** by incorporating **road closure interaction variables, or using a geospatial modeling** approach

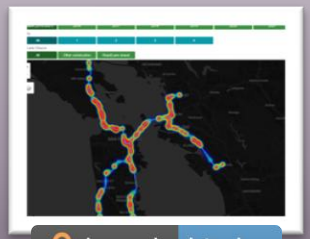
Refference Weblinks

- Contain all the codes (*.py, *.ipynb)
- Python environment file
- Presentation slide deck

https://bit.ly/4eqIn7v

- Hosted the GitHub notebook online with binder

https://bit.ly/44oM7l1



- Hosted the interactive python tools for assessing construction impact on Bay Area Routes

https://bit.ly/4l73xtV

Let's Discuss:
Your Questions - Our Path Forward