Michael S. Emanuel · Dylan Randle
Masters of Data Science

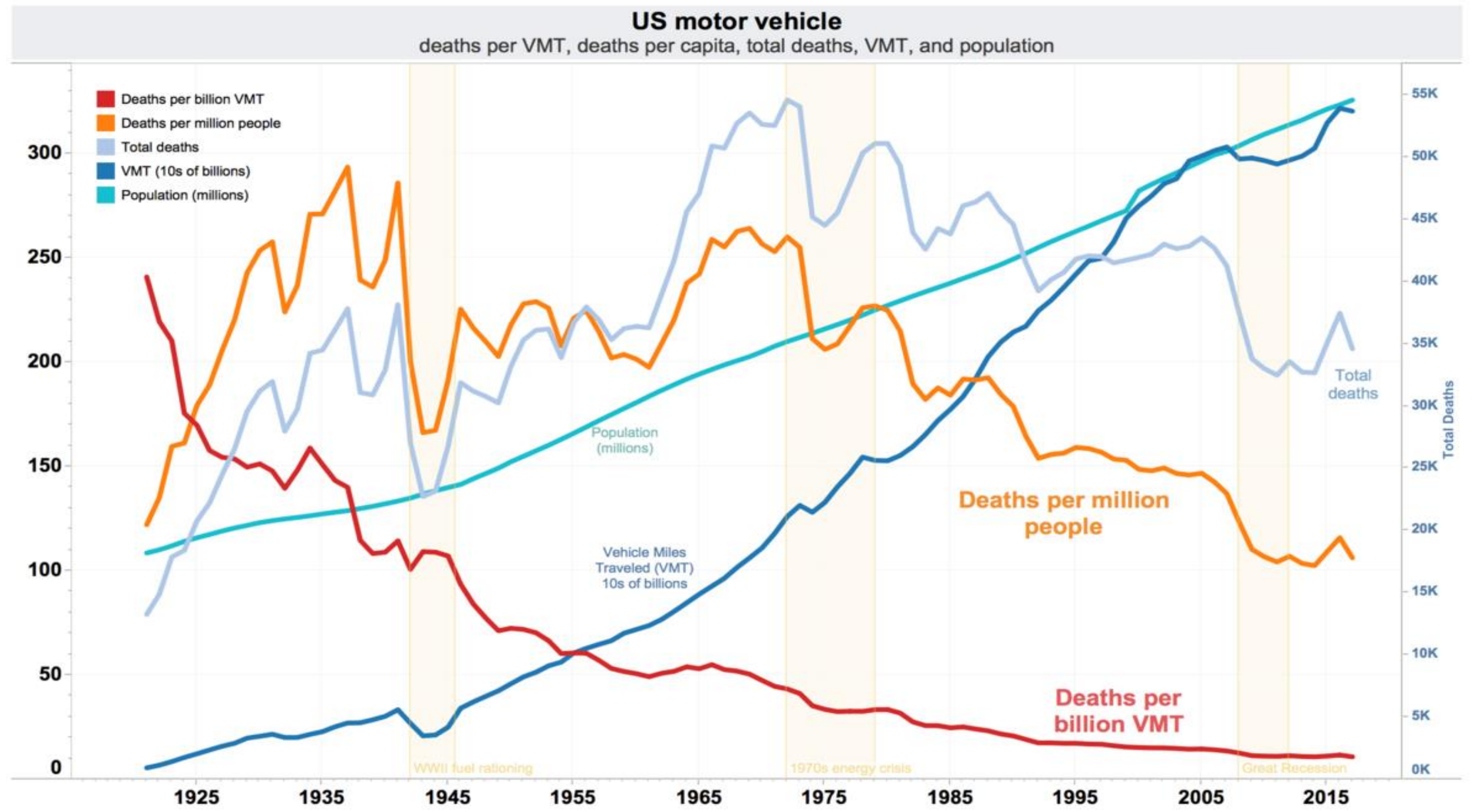
Safe Autonomous Vehicles

Technical and Non-Technical Solutions

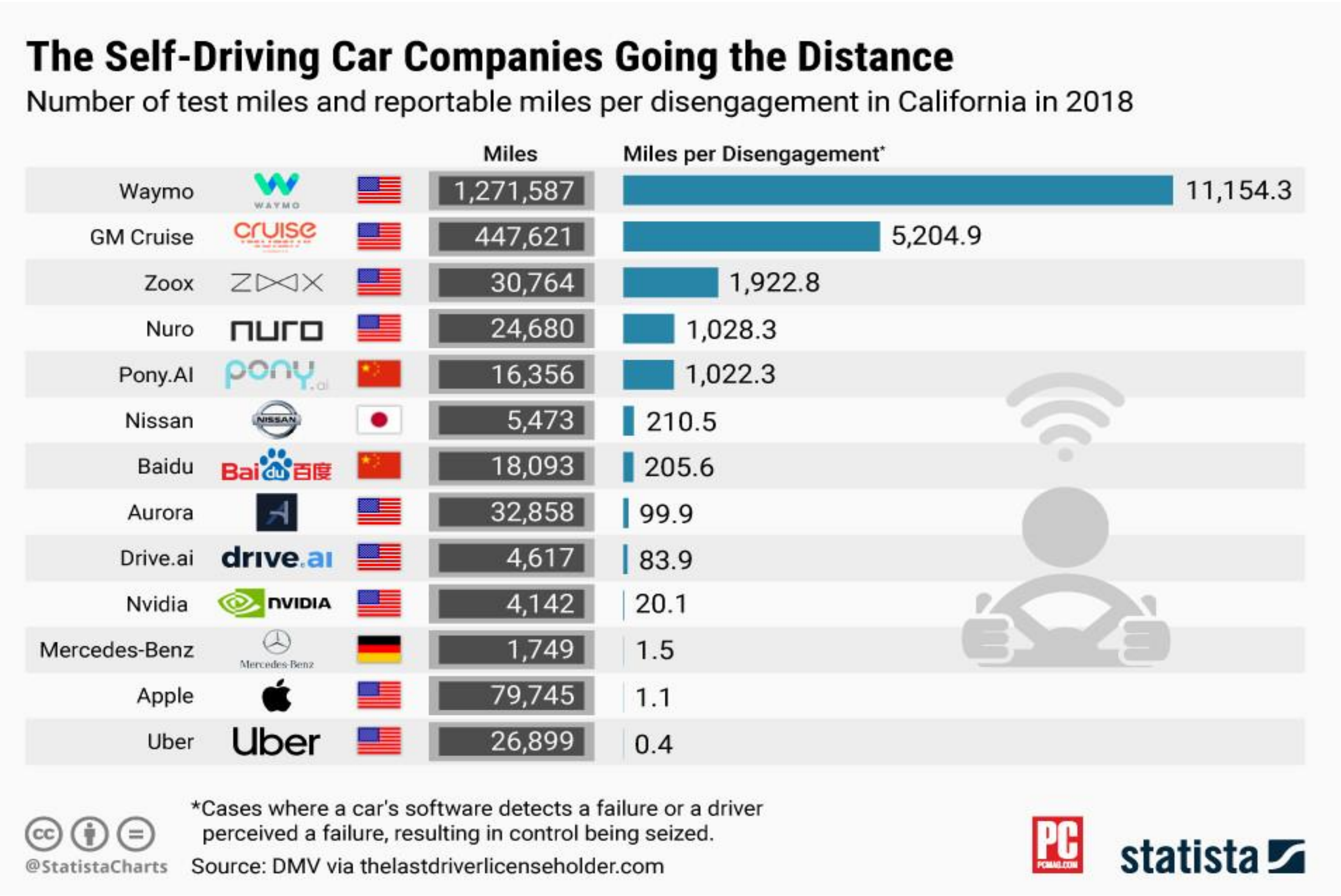
IACS AC 221
Critical Thinking in Data Science

Traffic is Deadly

Traffic accidents kill 37,000 Americans and 1.25 million people globally every year.
Forecasters estimate that AVs will save 585,000 lives 2035-2045 and generate $7 trillion revenue in 2050.



Waymo Leads the Pack & Improves Fast



The Self-Driving Car Companies Going the Distance

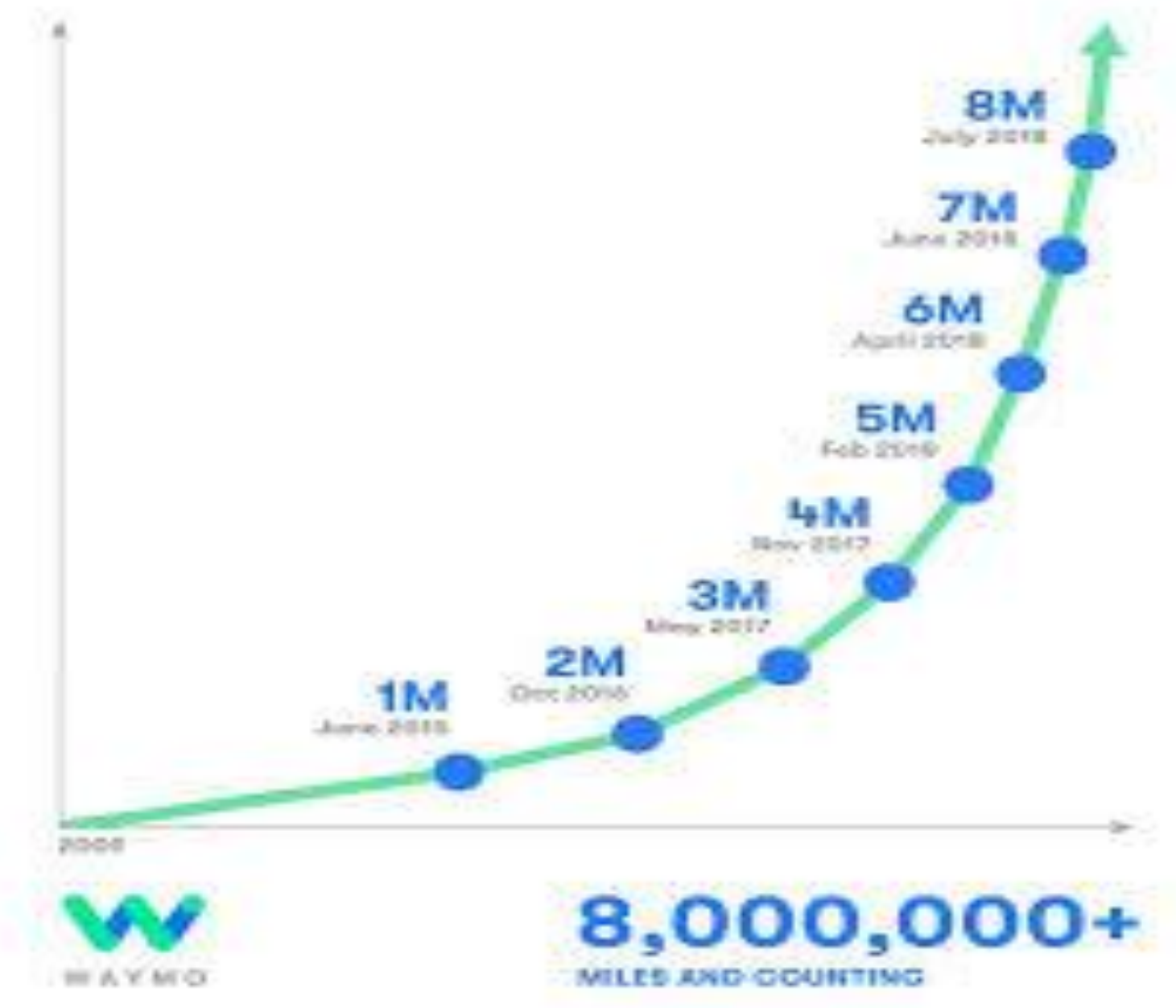
Number of test miles and reportable miles per disengagement in California in 2018

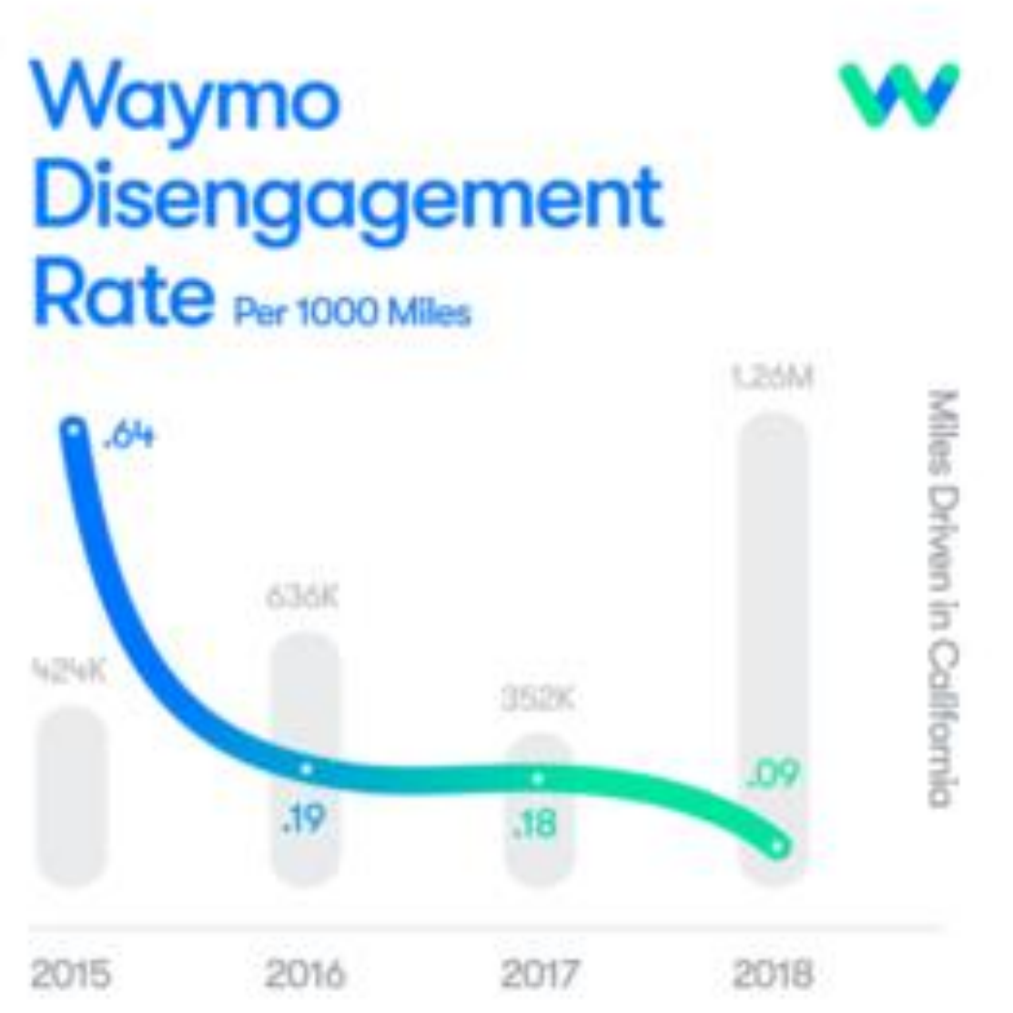
	Miles	Miles per Disengagement*
Waymo	1,271,587	11,154.3
GM Cruise	447,621	5,204.9
Zoox	30,764	1,922.8
Nuro	24,680	1,028.3
Pony.AI	16,356	1,022.3
Nissan	5,473	210.5
Baidu	18,093	205.6
Aurora	32,858	99.9
Drive.ai	4,617	83.9
Nvidia	4,142	20.1
Mercedes-Benz	1,749	1.5
Apple	79,745	1.1
Uber	26,899	0.4

*Cases where a car's software detects a failure or a driver perceived a failure, resulting in control being seized.

Source: DMV via thelastdriverlicenseholder.com

@StatistaCharts statista





Simulations Need Data

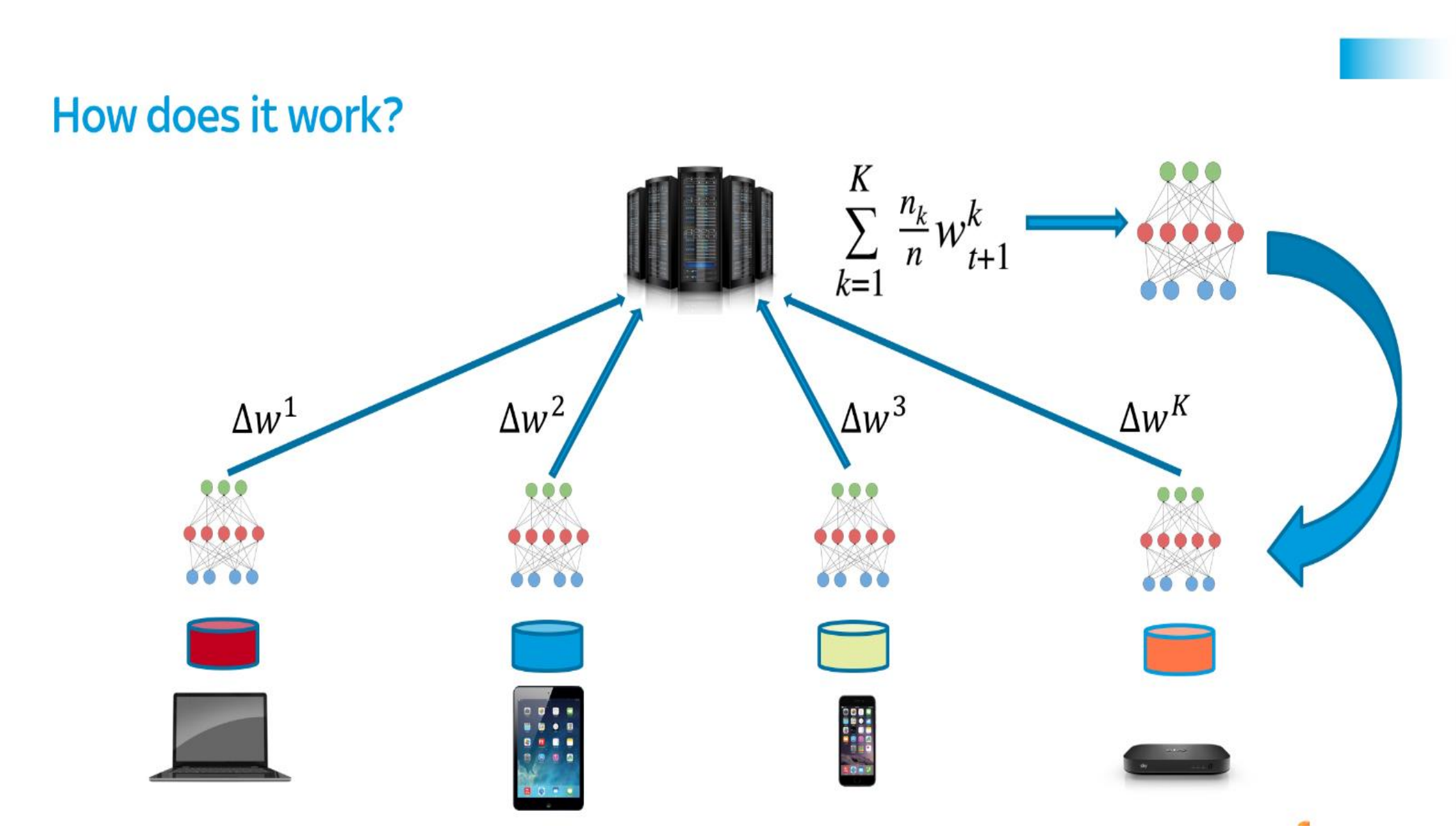
Simulation more important than road testing, which mainly validates the . Waymo has driven 10 *billion* simulator vs. 10 *million* real miles.
To “tame the tail” of rare driving events, the simulators need massive data.

Big Data = Human Drivers

Americans drive 8.8 billion miles a day. But driving data is too sensitive to share: accidents, traffic violations, legal liability, private itineraries.
Our proposal: Federated Learning to train neural nets while keeping driving data private.

Federated Learning & Secure MPC

A person can learn to drive without remembering everything they’ve seen on the road. So can a neural network!



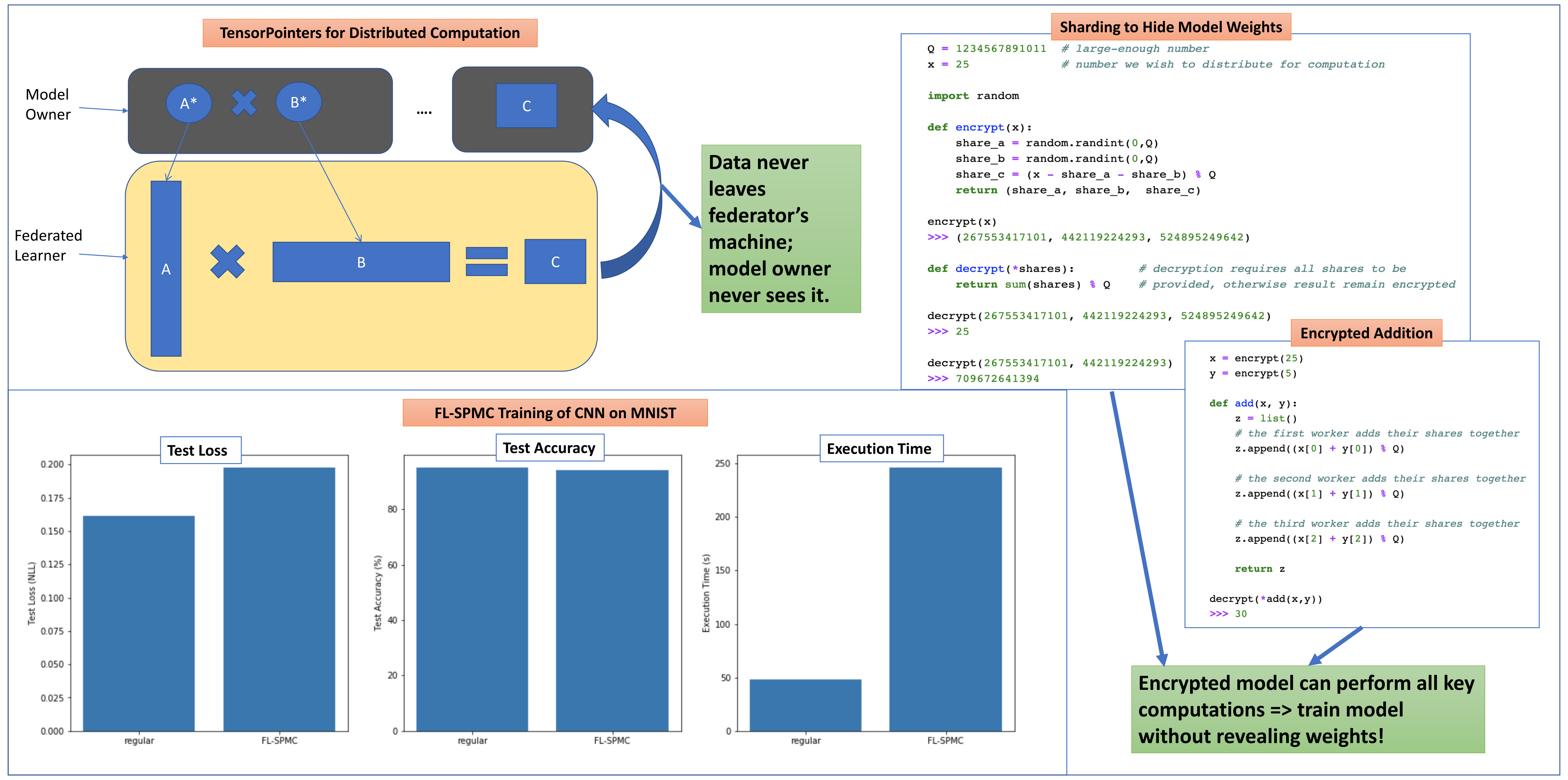
Uber Culture & Anthony Levandowski

“Safety can never be your No. 1 concern.”

“I’m pissed we didn’t have the first death.”

Anthony Levandowski led autonomous vehicle development at Uber before he was fired. Safety was never a top priority for him.
But company culture is about more than one person. When he was at Google, he quietly disabled a feature forbidding dangerous routes-and injured a colleague in a crash-before being overruled. At Uber, his emphasis on development speed over safety was unchallenged.

TensorPointers, Encrypted Programs, and Performance



Data never leaves federator’s machine; model owner never sees it.

Sharding to Hide Model Weights

```
Q = 1234567891011   # large-enough number
x = 25              # number we wish to distribute for computation

import random

def encrypt(x):
    share_a = random.randint(0,Q)
    share_b = random.randint(0,Q)
    share_c = (x - share_a - share_b) % Q
    return (share_a, share_b,  share_c)

encrypt(x)
>>> (267553417101, 442119224293, 524895249642)

def decrypt(*shares):           # decryption requires all shares to be
    return sum(shares) % Q      # provided, otherwise result remain encrypted

decrypt(267553417101, 442119224293, 524895249642)
>>> 25

decrypt(267553417101, 442119224293)
>>> 709672641394
```

Encrypted Addition

```
x = encrypt(25)
y = encrypt(5)

def add(x, y):
    z = list()
    # the first worker adds their shares together
    z.append((x[0] + y[0]) % Q)

    # the second worker adds their shares together
    z.append((x[1] + y[1]) % Q)

    # the third worker adds their shares together
    z.append((x[2] + y[2]) % Q)

    return z

decrypt(*add(x,y))
>>> 30
```

Encrypted model can perform all key computations => train model without revealing weights!

Conclusion: FL + Secure MPC is a Viable Way to Train AVs

We used a state of the art machine learning library, PySyft, to train a neural network on the classical computer vision task: MNIST digit classification. These results show that it should be feasible for a fleet of private vehicles controlled by people to train neural networks on key driving tasks such as image classification and scene segmentation while preserving the privacy of the underlying driving data of the human operators and the intellectual property of the model owners. However, there is a computational cost; communication between the federated learners and the model owner is costly.

Elayne Herzberg: First Pedestrian Killed by an AV

Elayne Herzberg was struck and killed by an Uber AV on March 18, 2018. She was the first victim of her kind. Herzberg was crossing outside of the crosswalk late at night. She first appeared on sensors 378 feet away, but the software failed catastrophically, classifying her as an “unknown object”, then as a bicycle, finally as a person. The system only called for emergency braking 76 feet before impact. But it didn’t brake because Uber engineers had disabled the feature; it made the ride too jerky. Instead it asked the safety driver to brake, but she was watching “The View” on her phone and didn’t brake until one second after impact.

Our Suggested Approach

We suggest some guidelines based on our values and AC 221 content:

- AVs should not be tested without safety drivers until they are **roughly as safe** as human operated vehicles
- AVs should not be widely deployed until they **meaningfully safer** than human operated vehicles, probably 2 to 10 times safer.
- People **hold machines to a higher safety standard** than human drivers. Right or wrong, this is unlikely to change.
- Regardless of comparisons to human operators, AVs should be engineered to be **as safe as reasonably possible** before deployments.
- **California’s disclosure regulations** on AV testing are working well
- Regulators may want to **require insurance for AVs** that is similar to insurance for human operators.

School of Engineering and Applied Sciences • Institute for Applied Computational Science