Agile Data Science

Elaine McVey
TransLoc

Marginality

Mystery

Misalignment

Marginality:

Perception that we are tangential to the core work of the organization

Mystery:

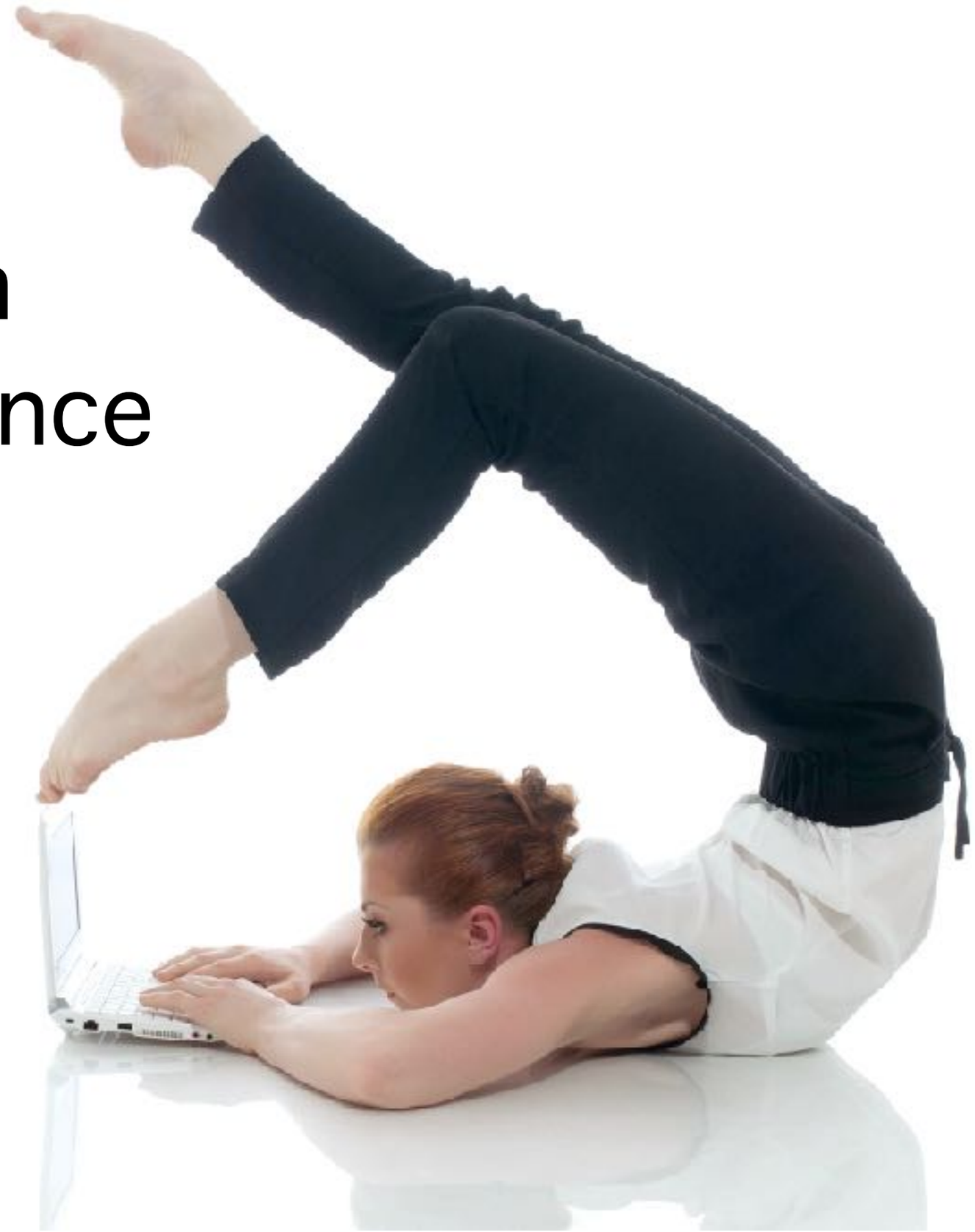
People struggle to really understand what we do

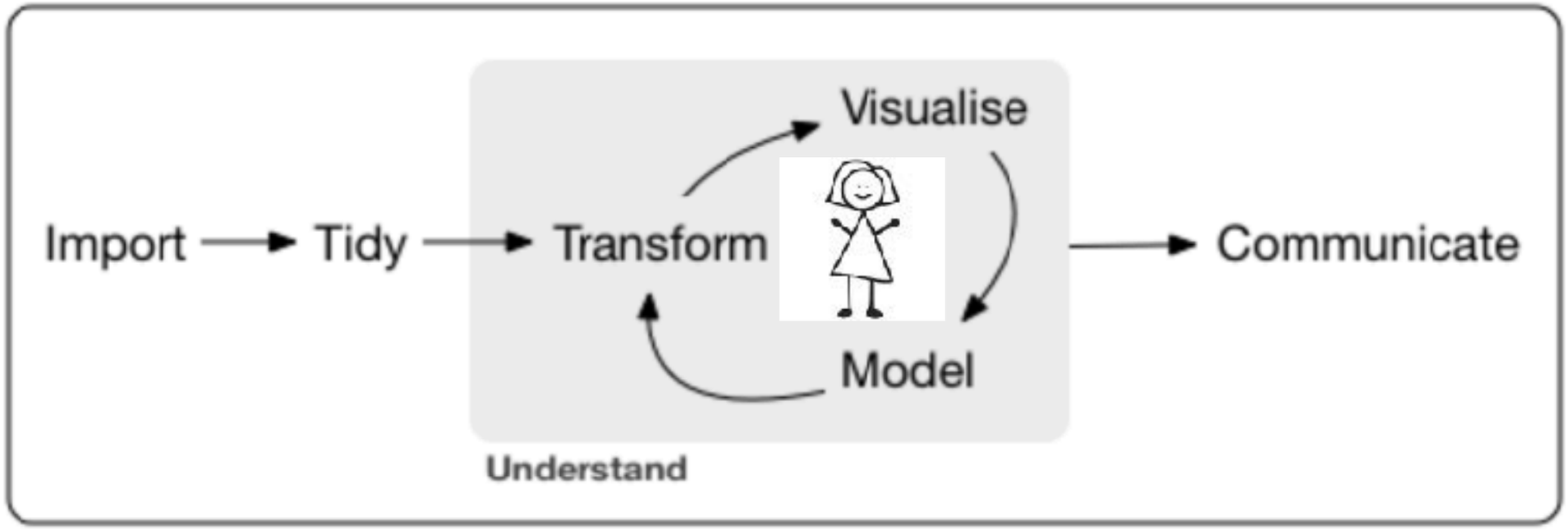
Misalignment:

Poor placement within the organization

Agile:

Change the conversation about data science





R for Data Science, Grolemund/Wickham

Agile:

Iterative process focused on delivering maximum value to the end user

Marginality

Mystery

Misalignment

#1: User Stories

Develop a shared understanding

We will simulate ride requests and algorithmic vehicle-trip assignments to produce a service schedule and metrics

As a transit planner, **I need to** understand how a microtransit system will perform in my region, **so that I can** assign the appropriate number of vehicles at launch

Marginality

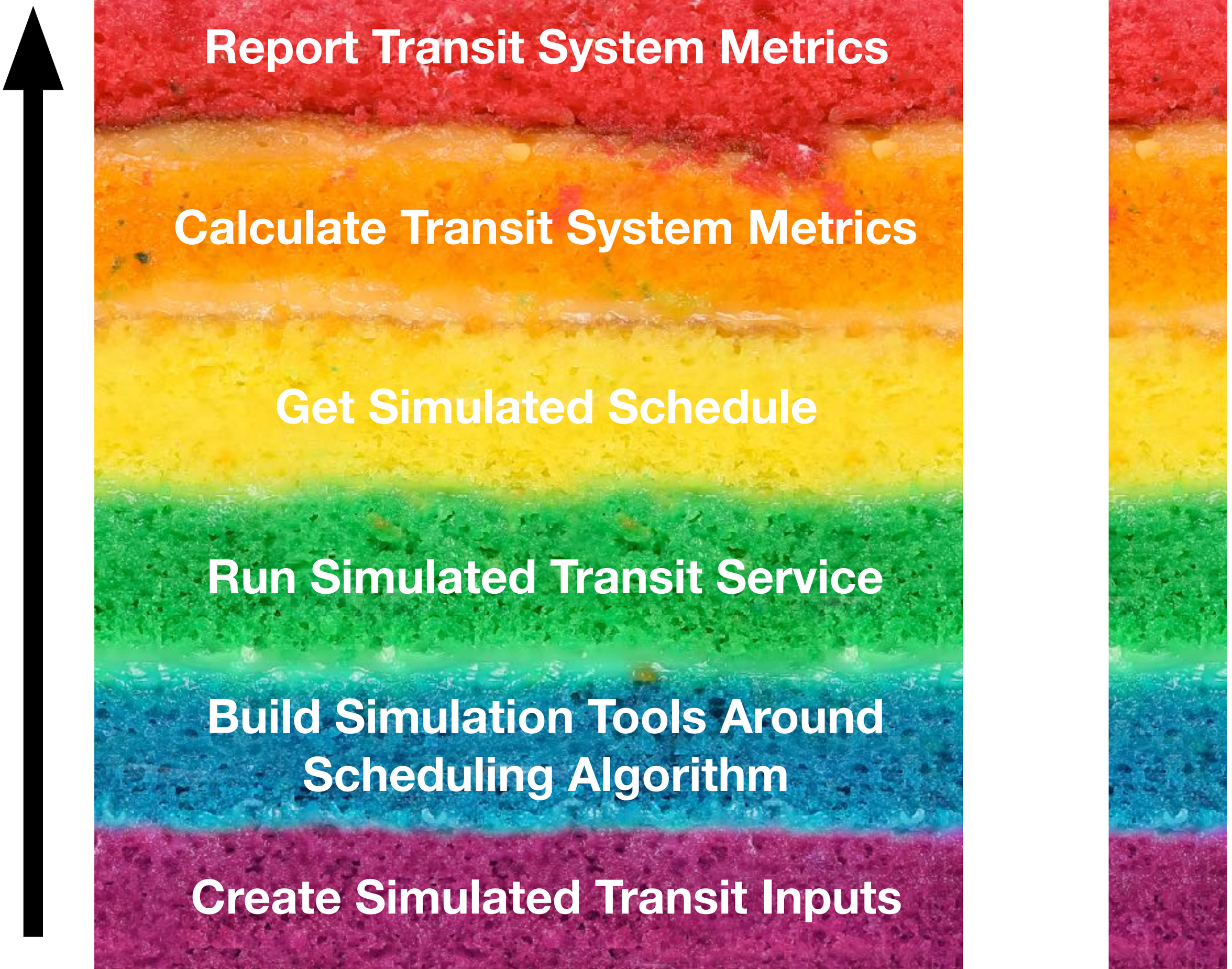
Shared
Value
Articulation

Mystery

Misalignment

#2: Thinnest Vertical Slice

Maximize the work not done


Report Transit System Metrics
Calculate Transit System Metrics
Get Simulated Schedule
Run Simulated Transit Service
Build Simulation Tools Around Scheduling Algorithm
Create Simulated Transit Inputs

Vehicles	Rides	Wait Time (min)	% Ride Sharing
10	300	17	36
15	300	11	21

Download:

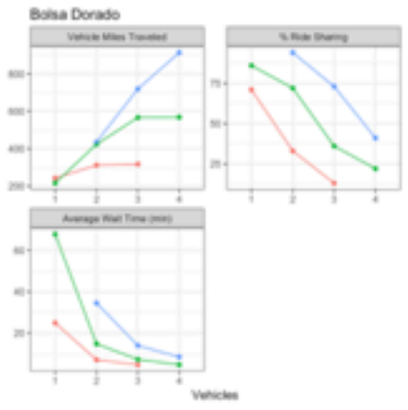
CSV Excel

Ride ID	Vehicle	Request time	Mins. until pickup	Pickup lat.	Pickup long.	Pickup time	Dropoff time	Mins. on vehicle	Drop
3	1	06:22	4	33.734	-117.999	06:26	06:34	8	33.
9	1	06:52	4	33.723	-118.041	06:56	07:10	14	33
11	1	06:56	6	33.748	-118.037	07:02	07:16	14	33.
17	1	07:12	10	33.73	-117.991	07:22	07:36	14	33
18	1	07:12	10	33.73	-117.991	07:22	07:36	14	33
19	1	07:12	10	33.73	-117.991	07:22	07:34	12	33
20	1	07:16	10	33.735	-117.999	07:26	07:36	10	33
21	1	07:16	10	33.735	-117.999	07:26	07:40	14	33
22	1	07:16							

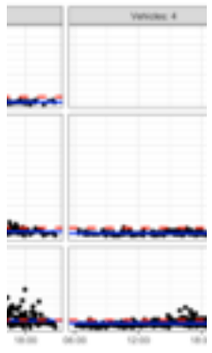
ributions

Service-Performance Tradeoffs

The graphs below show data from the above tables side-by-side to compare tradeoffs in service (wait times) and performance (vehicle miles traveled). Steep parts of the curves show where adding more vehicles can have a strong effect on the outcome.



for that simulation a



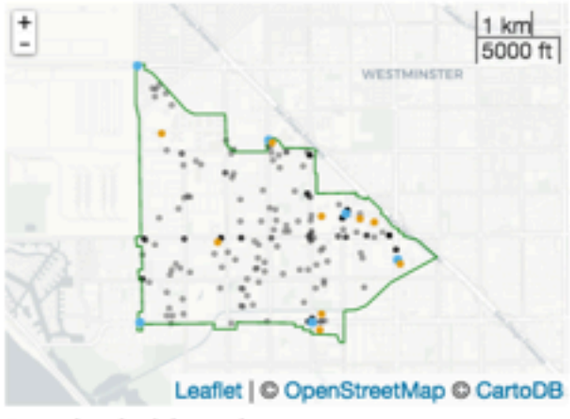
able below shows:
ed (% Ride Sharing),
vehicles when occu

des (% Time w/ 7+ R

Appendix

Trip Origins and Destinations

The map below shows trip origins (in black), the transit hubs (large blue markers) and other common destinations (large orange markers) for the simulation with 300 ride requests. Simulations with 100 and 200 rides were subsets of this data.



Example schedule results

Below are example results from the simulation with 300 ride requests per day and 3 vehicles.

Average Load	% Tim 7+ R
1.7	
1.1	
1.1	
4.5	
1.6	
1.2	
1.1	
3.1	
1.5	
1.2	

All

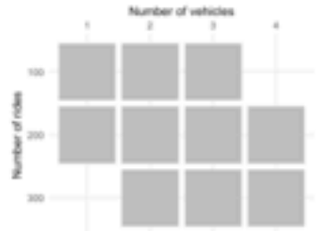
Service Design

Each simulation requires service parameters including number of vehicles, vehicle capacity, vehicle origin, time of service, and service type.

- Number of Vehicles: Simulations were run with combinations of 1, 2, 3, and 4 vehicles. We added the 4 vehicle scenarios for high ridership after examining the initial results.
- Vehicle Capacity: All vehicles were assumed to have a maximum of 9-passenger capacity.
- Vehicle Origin: At the beginning of the service period, all vehicles were assumed to be coming from 7301 Center Avenue, Huntington Beach, CA 92647.
- Time of Service: Ride requests were accepted for pickup times from 6:00 to 21:00.
- Service Type: Ride requests were simulated as on-demand, meaning that the ride would be scheduled for pickup as soon as possible after being requested.

Ride and Vehicle Combinations

We omitted the results with 300 rides and 1 vehicle because this combination had low quality results (only 52% of the rides were served and average wait times were over 1 hour) due to a backlog of requests and insufficient vehicle capacity.



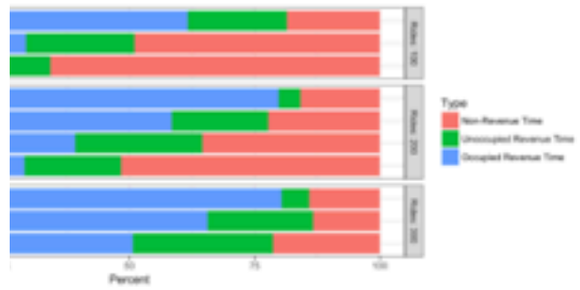
Results

The results for the daily service period in each case are below.

Operating Efficiency

ion is the number of rides served per vehicle hour.

a different number of rides, for comparisons across the s.



Time is idle time where vehicles are not in use.
venue Time is time when vehicles are en-route to pick

nue Time is time with riders (loading and driving).

Served (%)	Average Wait Time (min)	95%ile Wait Time (min)*	Average Ride Duration (min)	95%ile Ride Duration (min)*	Average Total Trip Time*
100	24.9	66.3	10	28.1	35
100	7	14.1	6.5	12.2	13.4
100	4.6	8.1	5.6	10	10.2
90	67.8	232.4	22.4	66.2	90.1
					23.8
					13.6
					10.7
					52.5
					23.8
					16

FM/LM | Interactive map

About this tool

Hide this information

Select Agency and Radius

Agency

Cerritos On Wheels

SCRTA

CDTA

Cecil Transit

Central Arkansas Transit Authority

Central Florida Regional Transit Authority

Centro

CGTSNET

Data sources

American Community Survey 5-year estimates 2010-2014

TransLoc stop locations as of 1/18/16

Public GTFS stop locations as of 3/15/16

Results

Bellflower

Artesia

Cerritos

Lakewood

La Palma

Bingham

Cerritos College

Gahr High School

Cerritos High School

Cerritos Regional County Park

MAPLE ST

WALNUT ST

PARK ST

CEDAR ST

166TH ST

183RD ST

195TH ST

CANDLEWOOD ST

DEL AMO BLVD

ARTESIA FWY

CA 91

I 605

NORWALK BLVD

PALO VERDE AVE

STUDEBAKER RD

GRIDLEY RD

FRESCA DR

Leaflet | Maps by Mapbox

Marginality

User Stories

Shared Value Articulation

Mystery

Vertical Slicing

Shared Value Validation

Misalignment

#3: Stakeholder Review

Collective resolution of competing priorities

1) What we did (and why)
2) What we propose to do next (and why)
3) Let's chat

DS Team Capacity = 9 units

Simulation Cohort = 4 units

Simulator Upgrades = 3 units

Pilot Analytics = 4 units

Microtransit Advisor = 2 units

Marginality

User Stories

Shared Value Articulation

Mystery

Vertical Slicing

Shared Value Validation

Misalignment

Stakeholder Reviews

Shared Value Prioritization

Agile requires discipline

datawoman.com

@eamcvey