Data Analysis & Exploration on Ride Hailing Data

Instructions

We want to know which cities are the healthiest, where there is room for improvement and how you might go about this. Use the attached spreadsheet for this task and produce a short presentation to walk through the reasoning that you use to answer the questions:

- How would you define the health of a city? What would be your key metric(s)?
- Using this definition which are the healthiest and least healthy cities and why?
- If you had limited resources to work on improving the health of cities, which cities would be your top priorities? Which metrics would you use to decide?
- For a high priority city, what do you think are the major health issues / opportunities? What types of
- operational and product improvements could we make to improve the marketplace?
- What other data sources would you want in an ideal world to help answer any of the questions above?

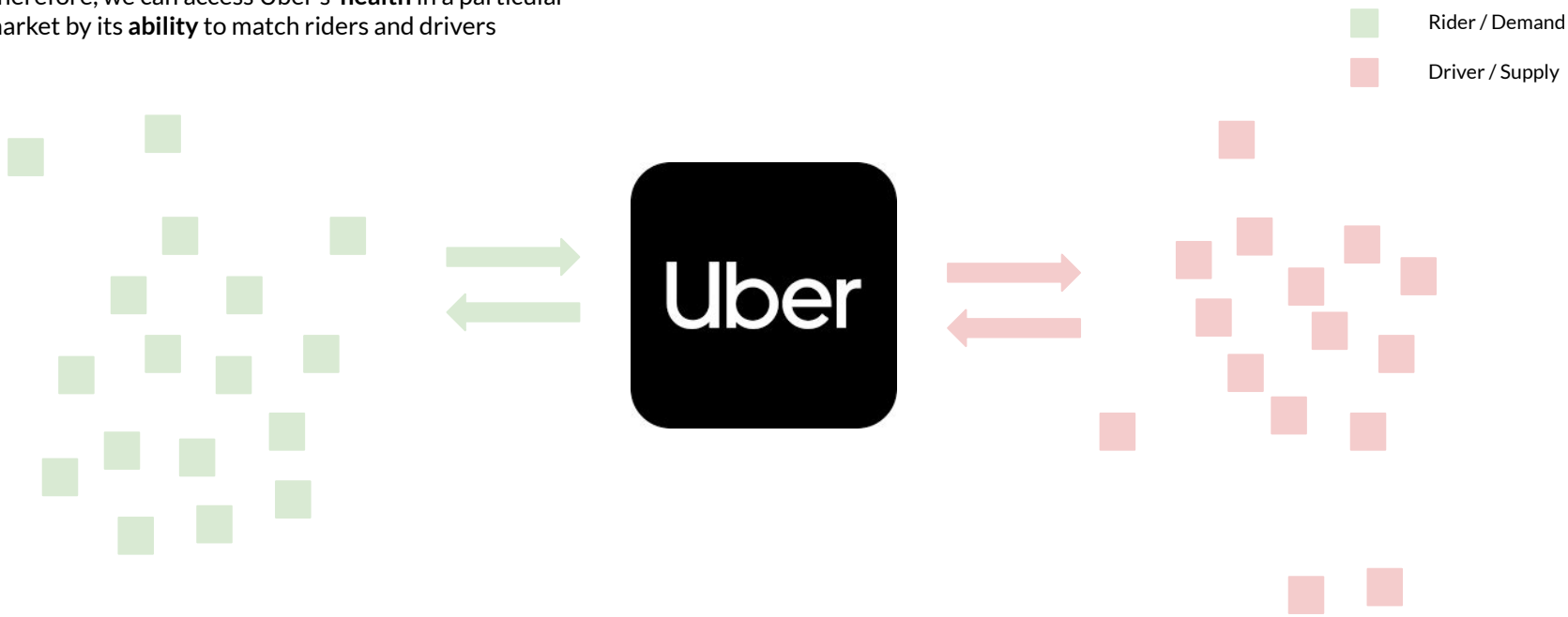
City Health

1. Understanding the Marketplace
2. City Health: Definition & Metrics
3. City Health: Calculation & Rank
4. City Health: Diagnosis & Prioritisation
5. Wirran Health: Diagnosis & Action
6. Other Data Source: Further Analysis

Understanding the Marketplace

Uber operates in a two-sided **marketplace** that connects riders with drivers.

Therefore, we can access Uber's **health** in a particular market by its **ability** to match riders and drivers

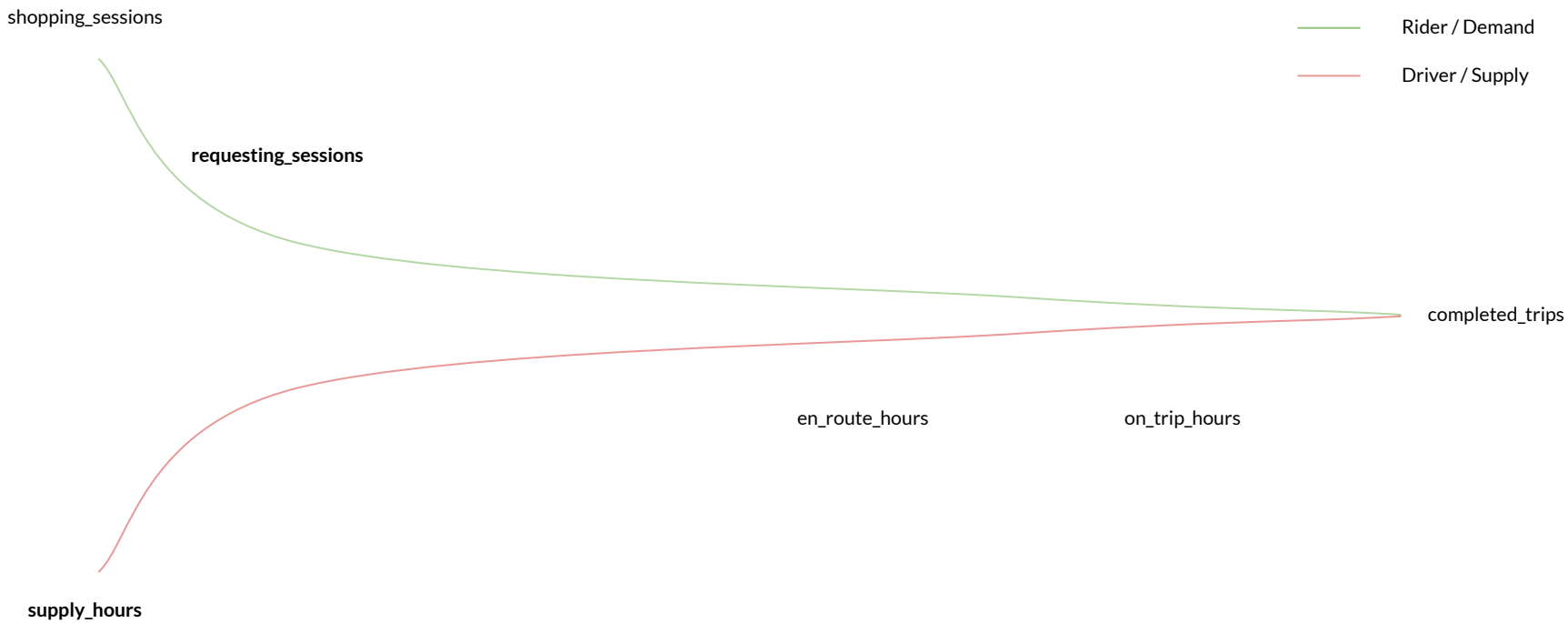


City Health: Definition & Metrics

Definition: the **efficiency** of a **sizeable market** that matches riders (demand) with drivers (supply)*

Sizeable market refers to volume of both demand and supply

Key Metrics: requesting_sessions (growth), supply_hours (growth)

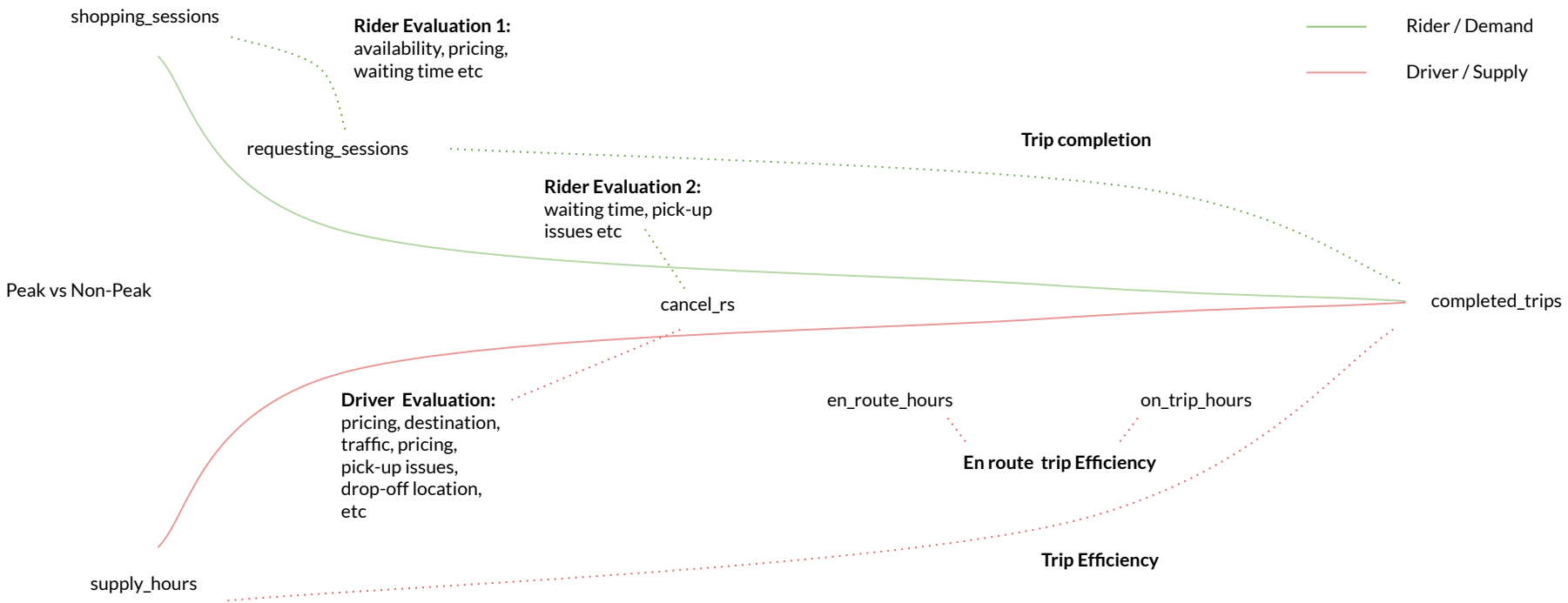


** metric calculation will be provided in python notebook*

City Health: Definition & Metrics

Definition: the **efficiency** of a **sizeable market** that matches riders (demand) with drivers (supply)

Efficiency refers to ability to match demand with supply

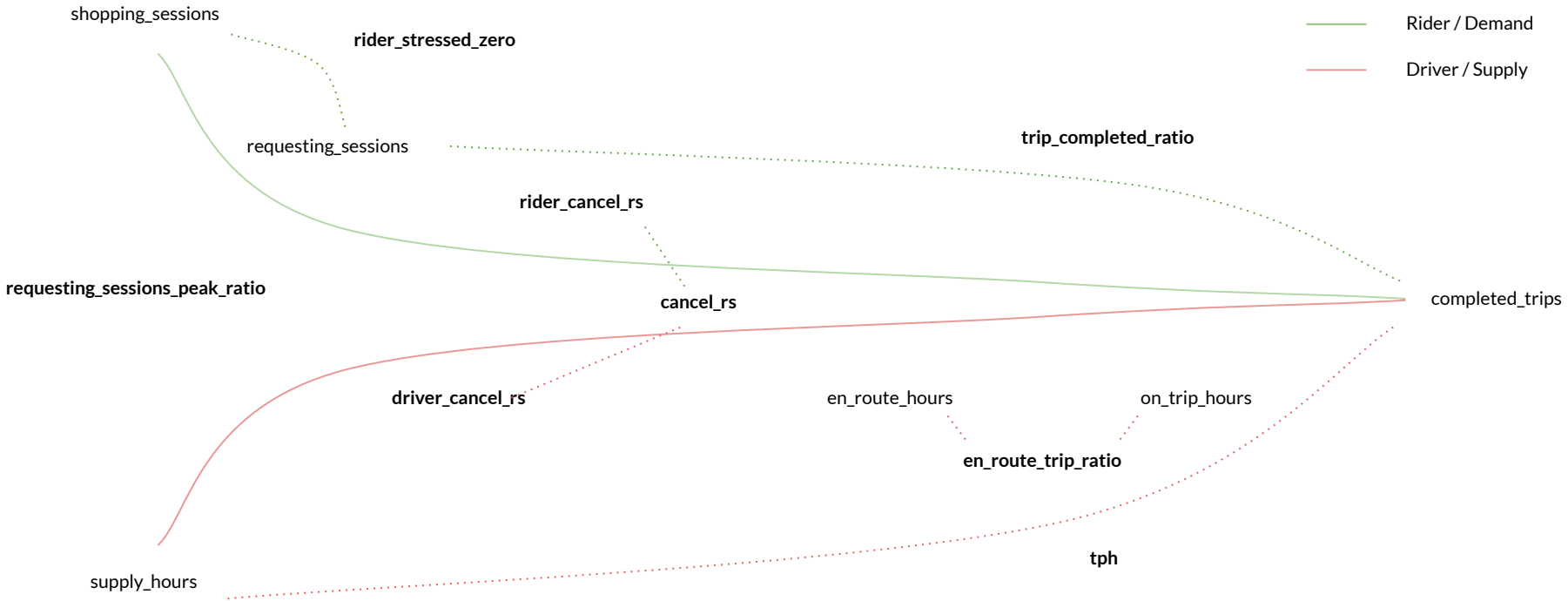


City Health: Definition & Metrics

Definition: the **efficiency** of a **sizeable market** that matches riders (demand) with drivers (supply)

Efficiency refers to ability to match demand with supply

Key Metrics: requesting_sessions_peak_ratio, rider_stressed_zero, rider_cancel_rs, trip_completed_ratio, cancel_rs, driver_cancel_rs, en_route_trip_ratio*, tph



** metric calculation will be provided in python notebook*

City Health: Calculation & Rank

Calculate each metric for every city, aggregate each metric score and rank overall city health

- 2 sides of marketplace
- 11 cities
- 12 metrics
- Complex riding process

- Calculate **size** metrics of the market (demand vs supply)
- Calculate Demand-supply matching **efficiency** metrics
- **Rank** cities based on each metric
- Weight*, aggregate rank scores and get **Overall City Rank**

Healthiest cities:
1. Allendale 2.Merian 3.Hillfield

Least Healthy cities:
11.Tapton 10.Wirran 9.Dunton

Good Bad

City Name	Size: Demand vs Supply				Efficiency							
	requesting_sessions	requesting_sessions_growth	supply_hours	supply_hours_growth	rider_stressed_zero	rider_cancel_rs	driver_cancel_rs	cancel_rs	en_route_trip_ratio	tph	trip_completed_ratio	requesting_sessions_peak_ratio
Allendale	75531	1.48%	28750	1.15%	41.36%	2.27%	0.88%	3.16%	52.15%	2.49	91.0%	21.0%
Brockpool	38403	1.34%	14133	0.91%	50.63%	3.73%	1.26%	4.99%	53.66%	2.39	84.9%	22.3%
Collinville	7248	1.74%	2568	1.54%	57.88%	4.55%	0.84%	5.38%	65.88%	2.46	83.9%	20.3%
Dunton	20934	1.38%	3546	2.20%	36.71%	10.10%	4.04%	14.13%	45.83%	0.74	66.1%	20.0%
Hillfield	43852	1.73%	12162	1.02%	56.85%	3.44%	0.91%	4.35%	54.80%	3.27	87.7%	20.6%
Lamberton	51493	1.21%	17201	1.02%	52.55%	3.83%	0.95%	4.78%	55.14%	2.69	84.7%	23.2%
Lorton	498852	0.90%	216498	0.61%	31.24%	2.38%	1.83%	4.22%	41.42%	2.15	90.8%	23.5%
Merian	245203	1.41%	96052	0.82%	38.76%	2.18%	1.31%	3.49%	45.11%	2.39	92.4%	22.2%
Northwood	115667	0.98%	42795	0.51%	56.67%	3.66%	1.37%	5.03%	46.38%	2.44	86.2%	22.5%
Tapton	8093	1.26%	3129	1.04%	62.40%	6.03%	1.15%	7.18%	57.59%	2.16	79.6%	23.4%
Wirran	45254	1.39%	15813	0.86%	59.35%	5.94%	1.81%	7.75%	48.81%	2.34	78.4%	22.3%

City Name	Size: Demand vs Supply				Efficiency								Rank	
	requesting_sessions	requesting_sessions_growth	supply_hours	supply_hours_growth	rider_stressed_zero	rider_cancel_rs	driver_cancel_rs	cancel_rs	en_route_trip_ratio	tph	trip_completed_ratio	requesting_sessions_peak_ratio	Weighted Total Rank Score	Overall City Rank
Allendale	4	3	4	3	4	2	2	1	6	3	2	4	52	1
Brockpool	8	7	7	7	5	6	6	6	7	6	6	7	107	8
Collinville	11	1	11	2	9	8	1	8	11	4	8	2	101	6
Dunton	9	6	9	1	2	11	11	11	3	11	11	1	111	9
Hillfield	7	2	8	6	8	4	3	4	8	1	4	3	81	3
Lamberton	5	9	5	5	6	7	4	5	9	2	7	9	97	5
Lorton	1	11	1	10	1	3	10	3	1	10	3	11	88	4
Merian	2	4	2	9	3	1	7	2	2	7	1	5	62	2
Northwood	3	10	3	11	7	5	8	7	4	5	5	8	103	7
Tapton	10	8	10	4	11	10	5	9	10	9	9	10	137	11
Wirran	6	5	6	8	10	9	9	10	5	8	10	6	117	10

** weight: 4 metrics for size vs 8 metrics for efficiency so to balance size and efficiency metrics, size metrics rank scores are doubled*

City Health: Diagnosis & Prioritisation

Limited resources, assuming it is our inabilities to **1) drive demand & supply in the short term** and **2) addressing all least healthiest cities**.

Therefore, the emphasis should be **improving efficiency**. In other words, **priorities would be cities with poor efficiencies**, which specifically lead to three cities - 11th Dunton, 10th Tapton, and 9th Wirran. Having said that, **Wirran** should be top priority as its **size** is much bigger

Good Bad

	Size: Demand vs Supply				Efficiency							
City Name	requesting_sessions	requesting_sessions_growth	supply_hours	supply_hours_growth	rider_stressed_zero	rider_cancel_rs	driver_cancel_rs	cancel_rs	en_route_trip_ratio	tph	trip_completed_ratio	requesting_sessions_peak_ratio
Allendale	75531	1.48%	28750	1.15%	41.36%	2.27%	0.88%	3.16%	52.15%	2.49	91.0%	21.0%
Brockpool	38403	1.34%	14133	0.91%	50.63%	3.73%	1.26%	4.99%	53.66%	2.39	84.9%	22.3%
Collinville	7248	1.74%	2568	1.54%	57.88%	4.55%	0.84%	5.38%	65.88%	2.46	83.9%	20.3%
Dunton	20934	1.38%	3546	2.20%	36.71%	10.10%	4.04%	14.13%	45.83%	0.74	66.1%	20.0%
Hillfield	43852	1.73%	12162	1.02%	56.85%	3.44%	0.91%	4.35%	54.80%	3.27	87.7%	20.6%
Lamberton	51493	1.21%	17201	1.02%	52.55%	3.83%	0.95%	4.78%	55.14%	2.69	84.7%	23.2%
Lorton	498852	0.90%	216498	0.61%	31.24%	2.38%	1.83%	4.22%	41.42%	2.15	90.8%	23.5%
Merian	245203	1.41%	96052	0.82%	38.76%	2.18%	1.31%	3.49%	45.11%	2.39	92.4%	22.2%
Northwood	115667	0.98%	42795	0.51%	56.67%	3.66%	1.37%	5.03%	46.38%	2.44	86.2%	22.5%
Tapton	8093	1.26%	3129	1.04%	62.40%	6.03%	1.15%	7.18%	57.59%	2.16	79.6%	23.4%
Wirran	45254	1.39%	15813	0.86%	59.35%	5.94%	1.81%	7.75%	48.81%	2.34	78.4%	22.3%

	Size: Demand vs Supply				Efficiency								Rank		
City Name	requesting_sessions	requesting_sessions_growth	supply_hours	supply_hours_growth	rider_stressed_zero	rider_cancel_rs	driver_cancel_rs	cancel_rs	en_route_trip_ratio	tph	trip_completed_ratio	requesting_sessions_peak_ratio	Weighted Total Rank Score	Overall City Rank	City Efficiency Rank
Allendale	4	3	4	3	4	2	2	1	6	3	2	4	52	1	1
Brockpool	8	7	7	7	5	6	6	6	7	6	6	7	107	8	5
Collinville	11	1	11	2	9	8	1	8	11	4	8	2	101	6	8
Dunton	9	6	9	1	2	11	11	11	3	11	11	1	111	9	9
Hillfield	7	2	8	6	8	4	3	4	8	1	4	3	81	3	3
Lamberton	5	9	5	5	6	7	4	5	9	2	7	9	97	5	5
Lorton	1	11	1	10	1	3	10	3	1	10	3	11	88	4	4
Merian	2	4	2	9	3	1	7	2	2	7	1	5	62	2	2
Northwood	3	10	3	11	7	5	8	7	4	5	5	8	103	7	5
Tapton	10	8	10	4	11	10	5	9	10	9	9	10	137	11	11
Wirran	6	5	6	8	10	9	9	10	5	8	10	6	117	10	10

** weight: 4 metrics for size vs 8 metrics for efficiency so to balance size and efficiency metrics, size metrics rank scores are doubled*

Wirran Health: Diagnosis & Action

Issues:

- **Stressed** and **zero** sessions account for 60% non requesting sessions - **limited availability and pricy**
- Both rider and driver **cancels** are high - possibly pick-up issues, long waiting time
- ultimately **low trip completion**

Opportunities:

- **Decent size** of demand and supply
- **Growing** weekly **requesting session**
- **Good en route trip ratio and tph -** efficient driving once committing a trip

Improvements:

- **Send** clear and frequent demand **forecast** to drivers
- **Increase estimated waiting time** to adjust **rider** expectation
- **Reduce commission** for **drivers** who have completed more trips

Good Bad

	Size: Demand vs Supply				Efficiency							
City Name	requesting_sessions	requesting_sessions_growth	supply_hours	supply_hours_growth	rider_stressed_zero	rider_cancel_rs	driver_cancel_rs	cancel_rs	en_route_trip_ratio	tph	trip_completed_ratio	requesting_sessions_peak_ratio
Allendale	75531	1.48%	28750	1.15%	41.36%	2.27%	0.88%	3.16%	52.15%	2.49	91.0%	21.0%
Brockpool	38403	1.34%	14133	0.91%	50.63%	3.73%	1.26%	4.99%	53.66%	2.39	84.9%	22.3%
Collinville	7248	1.74%	2568	1.54%	57.88%	4.55%	0.84%	5.38%	65.88%	2.46	83.9%	20.3%
Dunton	20934	1.38%	3546	2.20%	36.71%	10.10%	4.04%	14.13%	45.83%	0.74	66.1%	20.0%
Hillfield	43852	1.73%	12162	1.02%	56.85%	3.44%	0.91%	4.35%	54.80%	3.27	87.7%	20.6%
Lamberton	51493	1.21%	17201	1.02%	52.55%	3.83%	0.95%	4.78%	55.14%	2.69	84.7%	23.2%
Lorton	498852	0.90%	216498	0.61%	31.24%	2.38%	1.83%	4.22%	41.42%	2.15	90.8%	23.5%
Merian	245203	1.41%	96052	0.82%	38.76%	2.18%	1.31%	3.49%	45.11%	2.39	92.4%	22.2%
Northwood	115667	0.98%	42795	0.51%	56.67%	3.66%	1.37%	5.03%	46.38%	2.44	86.2%	22.5%
Tapton	8093	1.26%	3129	1.04%	62.40%	6.03%	1.15%	7.18%	57.59%	2.16	79.6%	23.4%
Wirran	45254	1.39%	15813	0.86%	59.35%	5.94%	1.81%	7.75%	48.81%	2.34	78.4%	22.3%

	Size: Demand vs Supply				Efficiency								Rank		
City Name	requesting_sessions	requesting_sessions_growth	supply_hours	supply_hours_growth	rider_stressed_zero	rider_cancel_rs	driver_cancel_rs	cancel_rs	en_route_trip_ratio	tph	trip_completed_ratio	requesting_sessions_peak_ratio	Weighted Total Rank Score	Overall City Rank	City Efficiency Rank
Allendale	4	3	4	3	4	2	2	1	6	3	2	4	52	1	1
Brockpool	8	7	7	7	5	6	6	6	7	6	6	7	107	8	5
Collinville	11	1	11	2	9	8	1	8	11	4	8	2	101	6	8
Dunton	9	6	9	1	2	11	11	11	3	11	11	1	111	9	9
Hillfield	7	2	8	6	8	4	3	4	8	1	4	3	81	3	3
Lamberton	5	9	5	5	6	7	4	5	9	2	7	9	97	5	5
Lorton	1	11	1	10	1	3	10	3	1	10	3	11	88	4	4
Merian	2	4	2	9	3	1	7	2	2	7	1	5	62	2	2
Northwood	3	10	3	11	7	5	8	7	4	5	5	8	103	7	5
Tapton	10	8	10	4	11	10	5	9	10	9	9	10	137	11	11
Wirran	6	5	6	8	10	9	9	10	5	8	10	6	117	10	10

** weight: 4 metrics for size vs 8 metrics for efficiency so to balance size and efficiency metrics, size metrics rank scores are doubled*

Other Data Source: Further Analysis

P&L information

- Customer lifetime value
- Cost per acquisition
- Pricing
- Commission
- Trip fair

Trip information

- Trip timestamp
- Starting/ending destination
- Weather
- Trip miles
- Trip hours/minutes

Driver / Rider information

- Number of trips
- Fair of trips
- Demographics
- Rating
- If other Uber service users