

Uber and Lyft Price Analysis

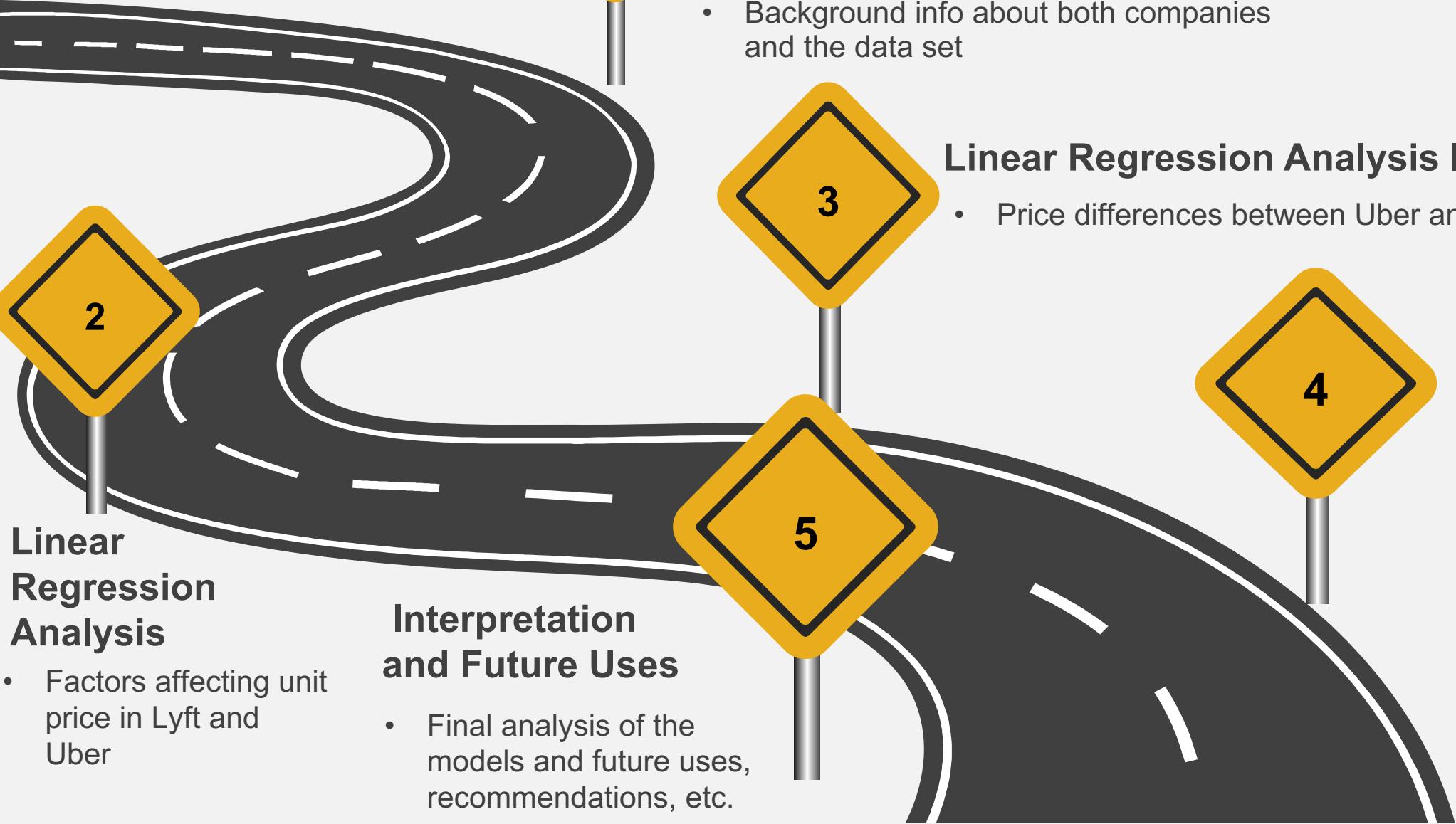
Analysis by:

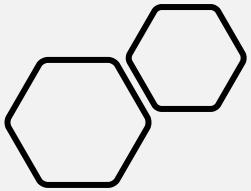
Taoyanran Sun, Cheng Ling, Dana Barclay, and Yun-Ting Sun





Roadmap





Company Backgrounds

Uber

- Founded in San Francisco, CA 2009
- Over 10 billion rides worldwide
- 69% market share

Lyft

- Founded in San Francisco, CA 2012
- Over 1 billion rides in the US and Canada
- 29% market share

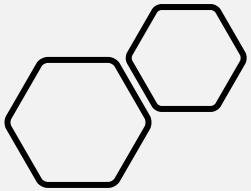
\$ Pricing Comparison for Uber and Lyft in Boston

UberX

- Minimum price - \$6.85
- Base fare - \$2.00
- Cost per mile - \$1.24
- Cost per minute - \$0.20

Lyft

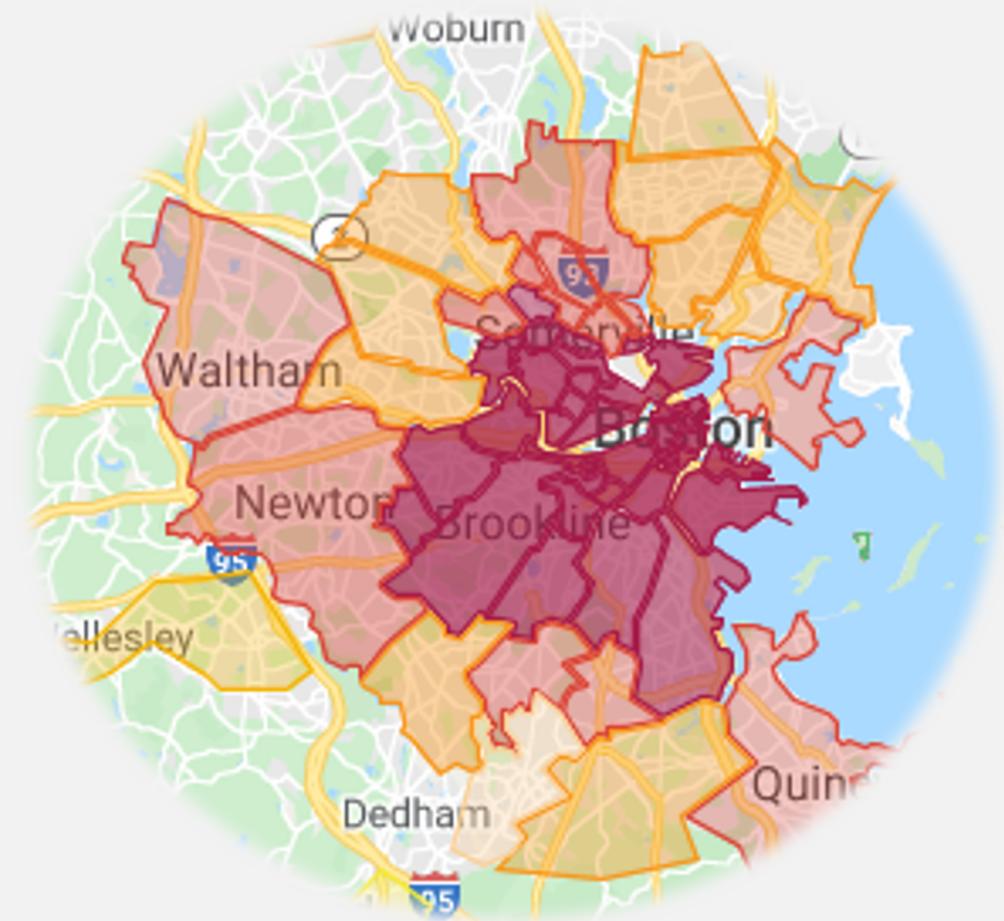
- Minimum price - \$5.00
- Base fare - \$2.10
- Cost per mile - \$0.88
- Cost per minute - \$0.36



Data Set Background

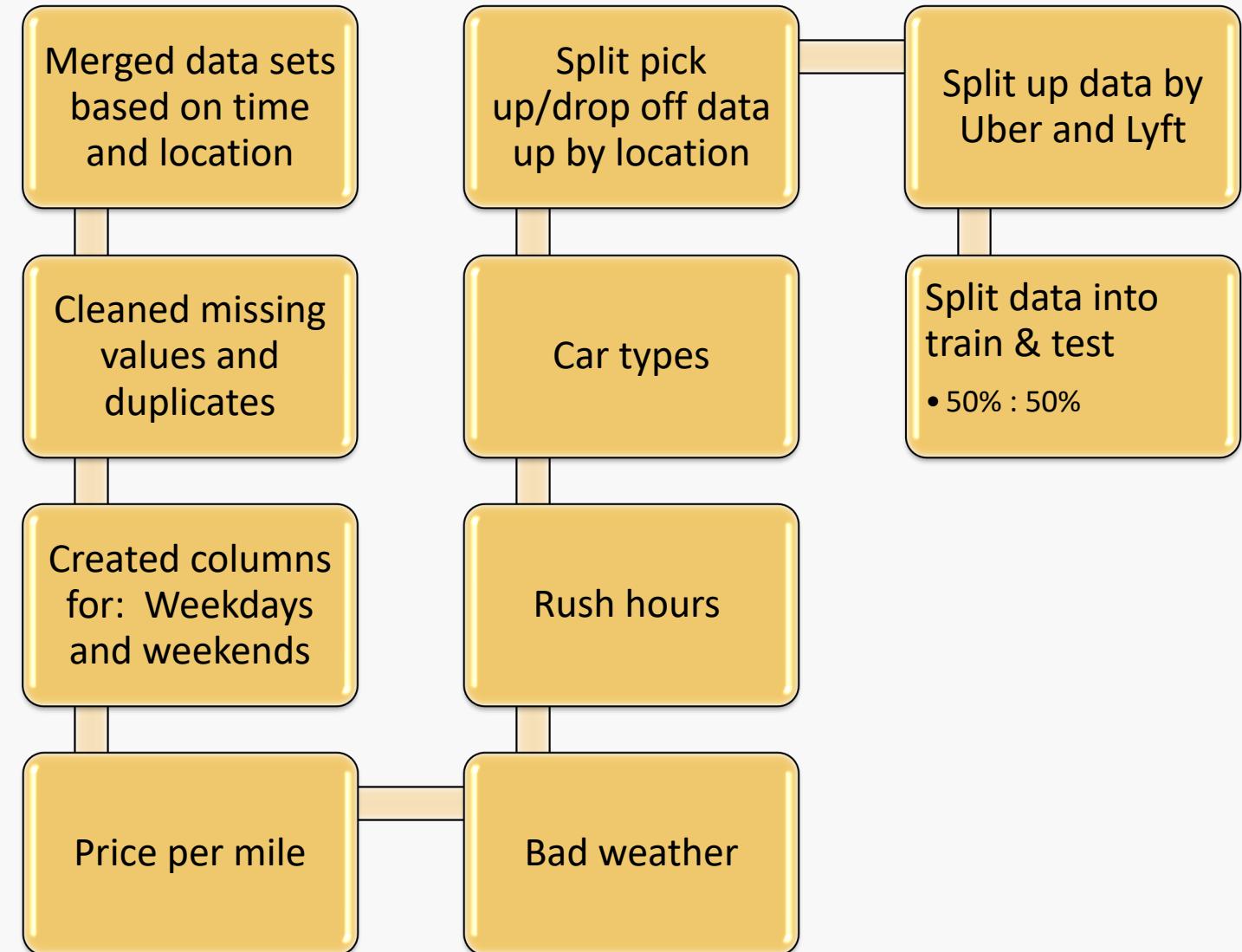
Uber & Lyft Cab Prices

- Data collected from "hot" locations in Boston Nov 26th – Dec 18th, 2018
- 632,403 observations, 18 variables total
- Contains cab data
 - Every 5 minutes
- Contains weather data
 - Every 1 hour



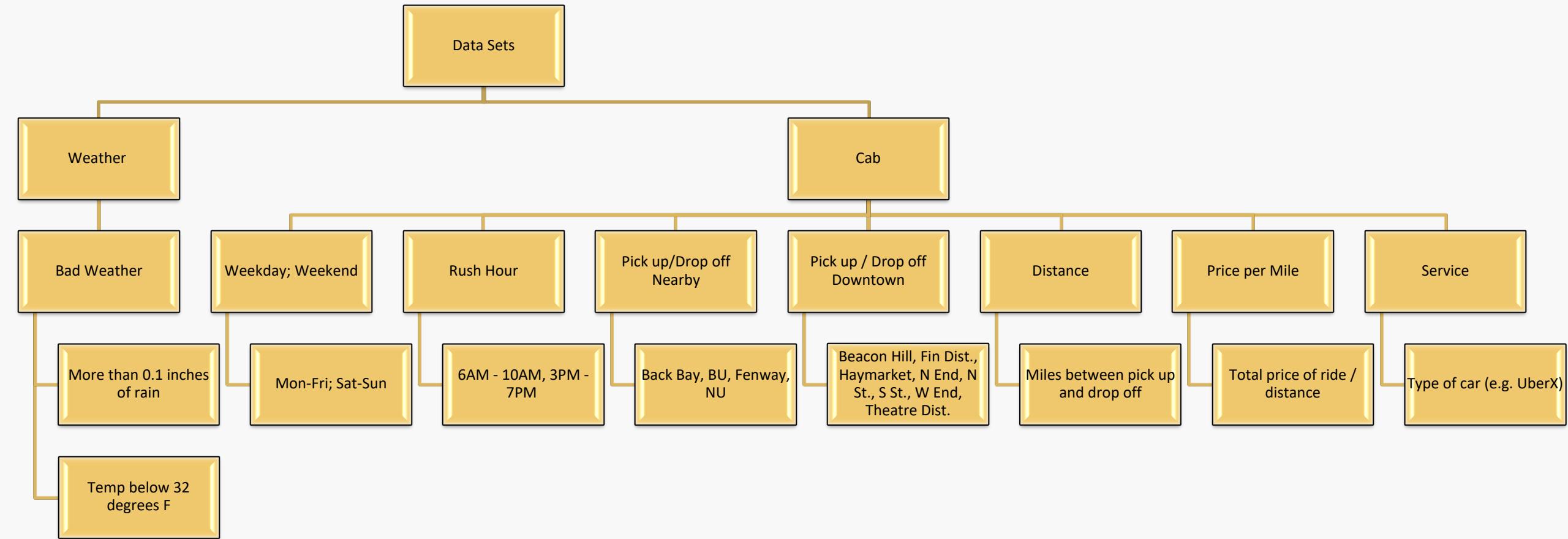


Data Preparation





Variable Definitions



Question 1

What are the factors that affect price per mile in Uber and Lyft?



Uber Linear Model and Regression Tree

```
lm(formula = log(price_per_mile) ~ I(distance^0.4) * service +  
  bad_weather + Isweekend + rush_hour + pickup + dropoff, data = train_u)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.74082	-0.22443	-0.05162	0.19394	3.03420

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.530e+00	4.058e-03	869.875	<2e-16 ***
I(distance^0.4)	-1.355e+00	3.155e-03	-429.393	<2e-16 ***
serviceUber Premium	9.680e-01	6.804e-03	142.270	<2e-16 ***
bad_weather1	-2.709e-03	1.973e-03	-1.373	0.170
Isweekend1	3.766e-05	1.629e-03	0.023	0.982
rush_hour1	-1.987e-03	1.898e-03	-1.047	0.295
pickupNearby	-4.261e-02	1.641e-03	-25.974	<2e-16 ***
dropoffNearby	-2.886e-02	1.610e-03	-17.919	<2e-16 ***
I(distance^0.4):serviceUber Premium	-8.419e-02	5.025e-03	-16.754	<2e-16 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2869 on 164561 degrees of freedom

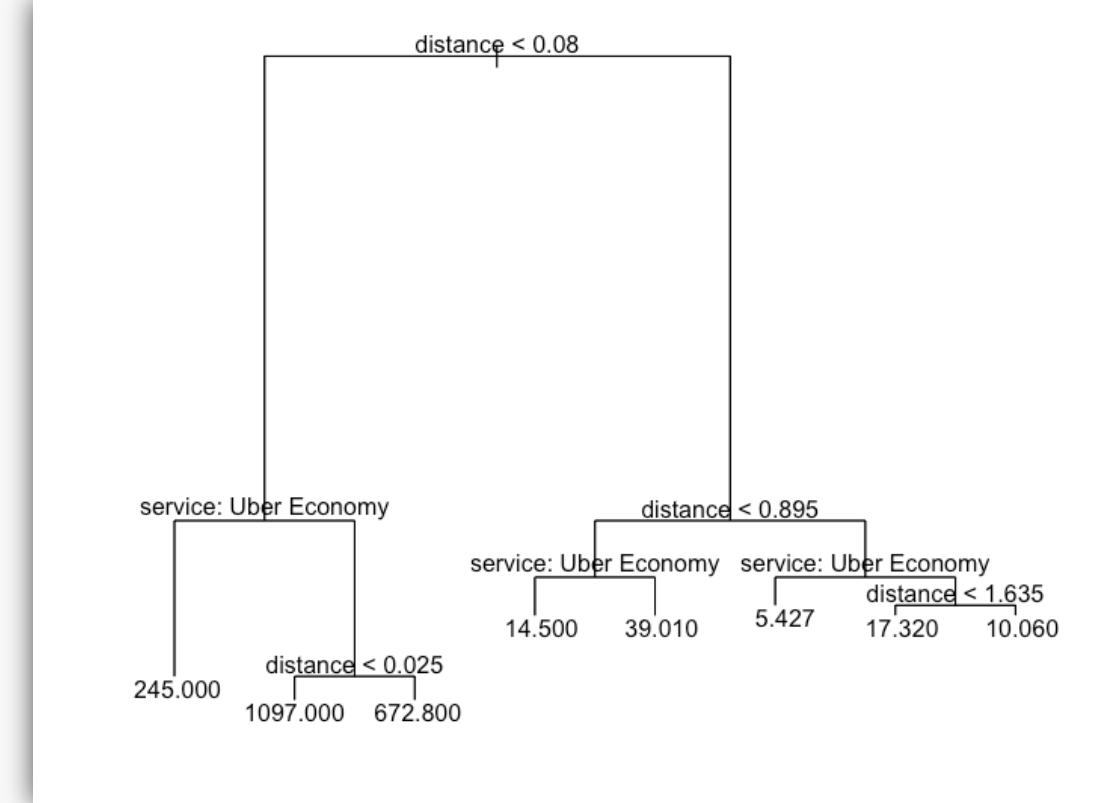
Multiple R-squared: 0.8059, Adjusted R-squared: 0.8059

F-statistic: 8.539e+04 on 8 and 164561 DF, p-value: < 2.2e-16

R² Adj = 0.8059

Test Rate = 0.25

MSE = 231.36



Best Size = 8

MSE = 35.59

Most Important Predictor: Distance

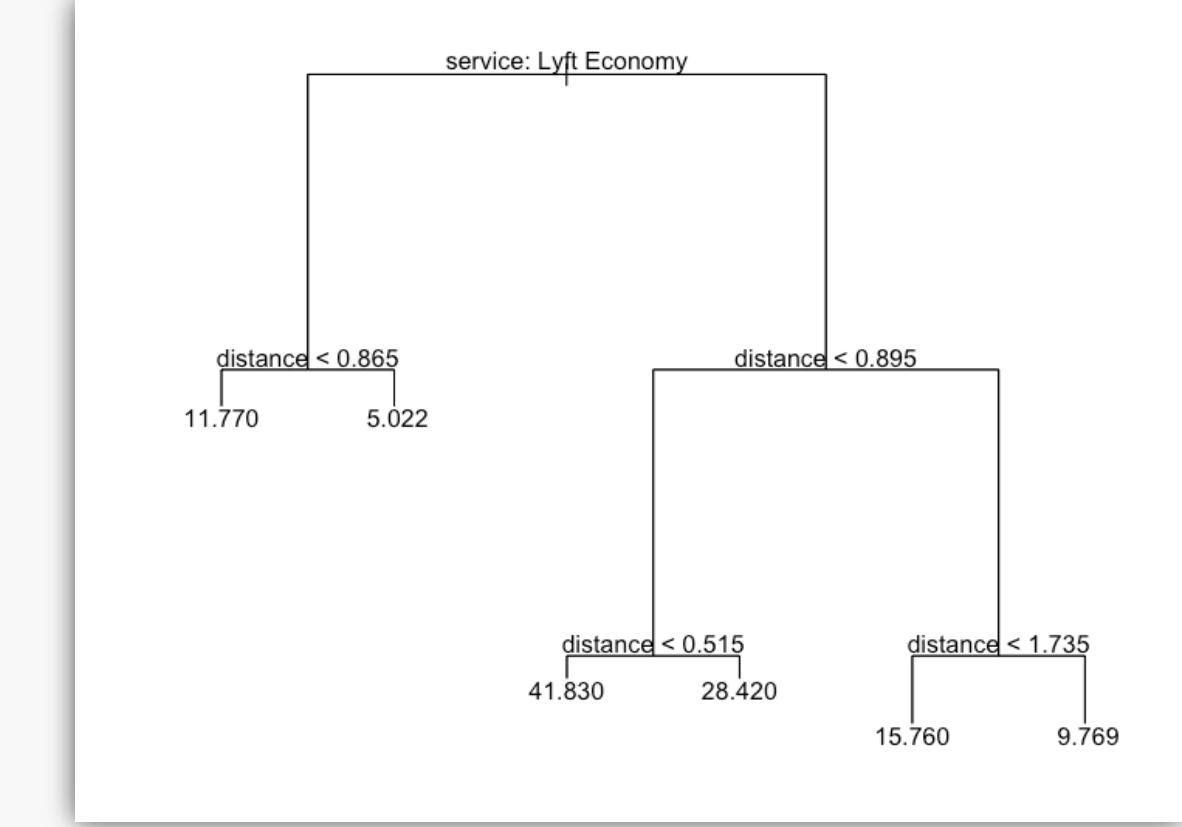
Lyft Linear Model and Regression Tree

```
Call:  
lm(formula = log(price_per_mile) ~ I(distance^0.4) + service +  
  bad_weather + Isweekend + rush_hour + pickup + dropoff +  
  surge_multiplier, data = train_l)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-1.47190 -0.25624 -0.02085  0.28655  0.99921  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 2.3369465 0.0084171 277.642 <2e-16 ***  
I(distance^0.4) -1.2075774 0.0037687 -320.423 <2e-16 ***  
serviceLyft Premium 0.9010964 0.0018623 483.850 <2e-16 ***  
bad_weather1 0.0001785 0.0025926 0.069 0.945  
Isweekend1 0.0008278 0.0021435 0.386 0.699  
rush_hour1 -0.0020629 0.0024983 -0.826 0.409  
pickupNearby -0.0478533 0.0021354 -22.409 <2e-16 ***  
dropoffNearby -0.0332124 0.0021788 -15.244 <2e-16 ***  
surge_multiplier 0.8471319 0.0068366 123.912 <2e-16 ***  
---  
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.3639 on 153042 degrees of freedom  
Multiple R-squared: 0.7203, Adjusted R-squared: 0.7203  
F-statistic: 4.926e+04 on 8 and 153042 DF, p-value: < 2.2e-16
```

R² Adj = 0.72

Test Rate = 0.66

MSE = 19.15



Best Size = 6

MSE = 18.59

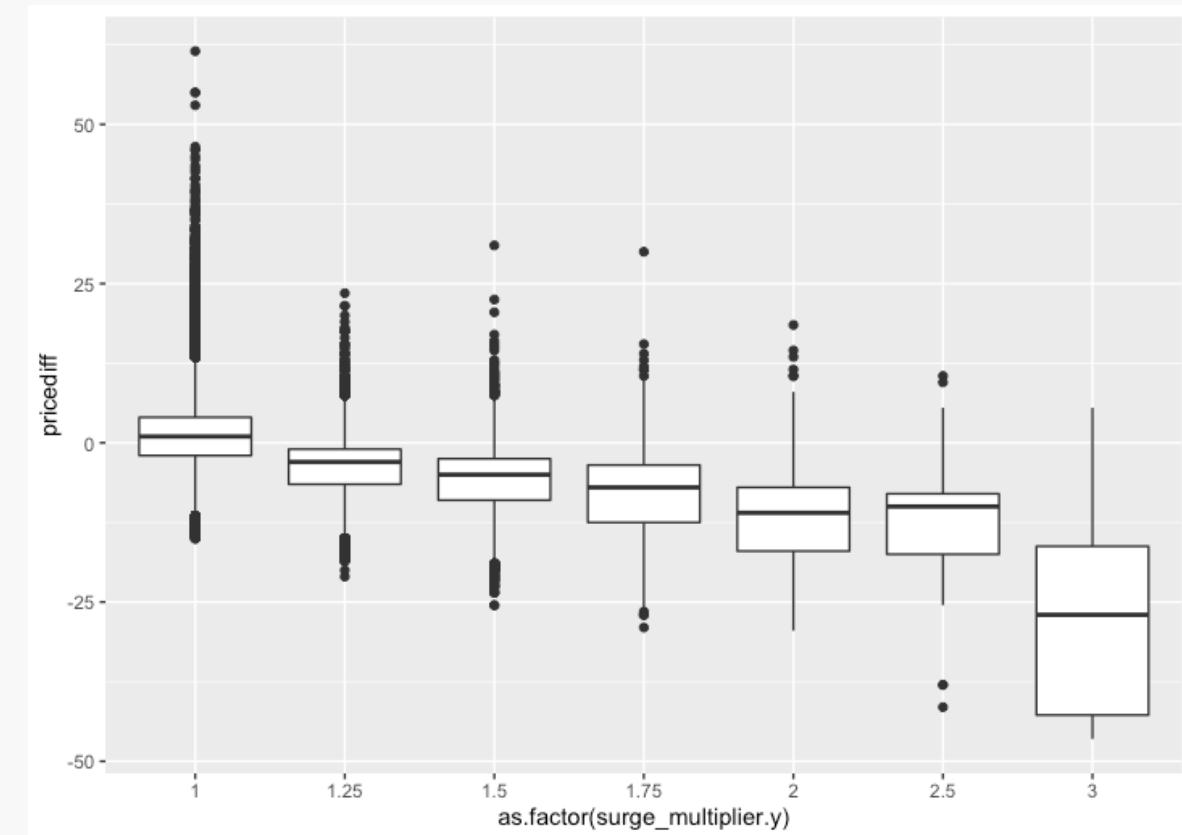
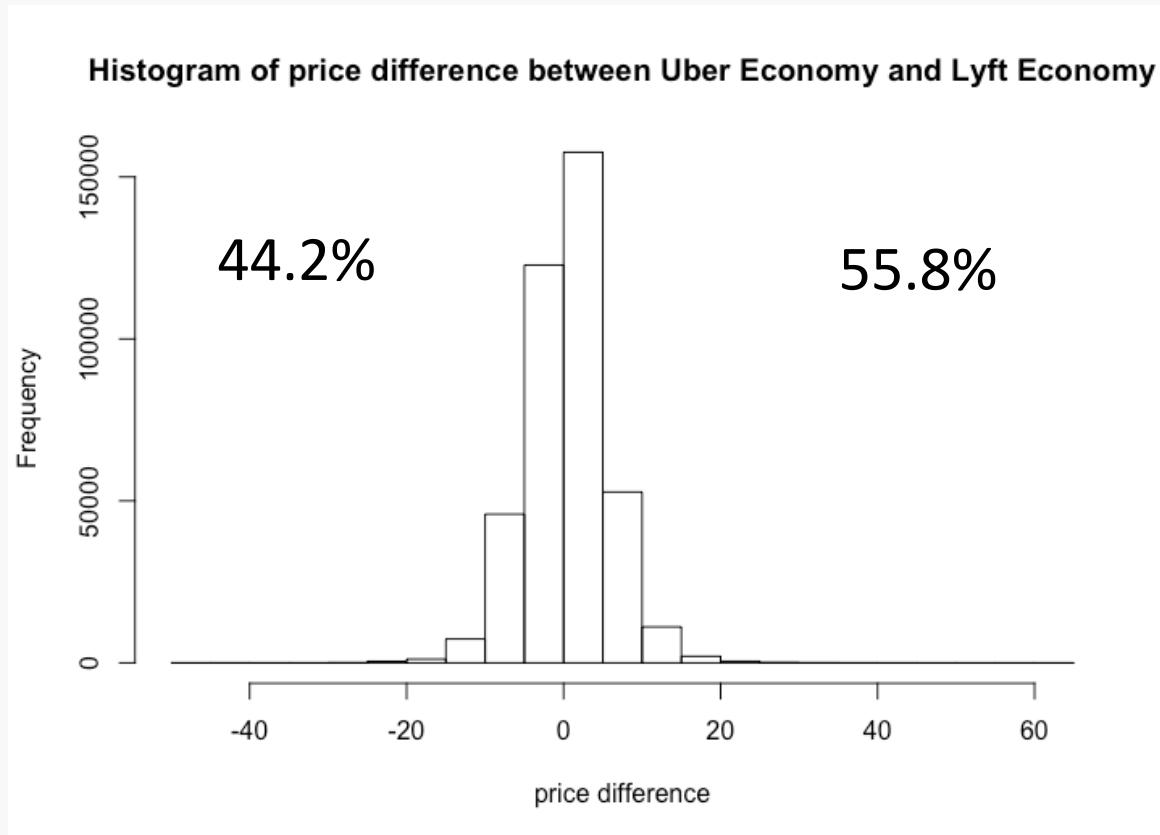
Most Important Predictor: Service(Lyft Economy)

Question 2

Is there a difference in
pricing between Uber and
Lyft?

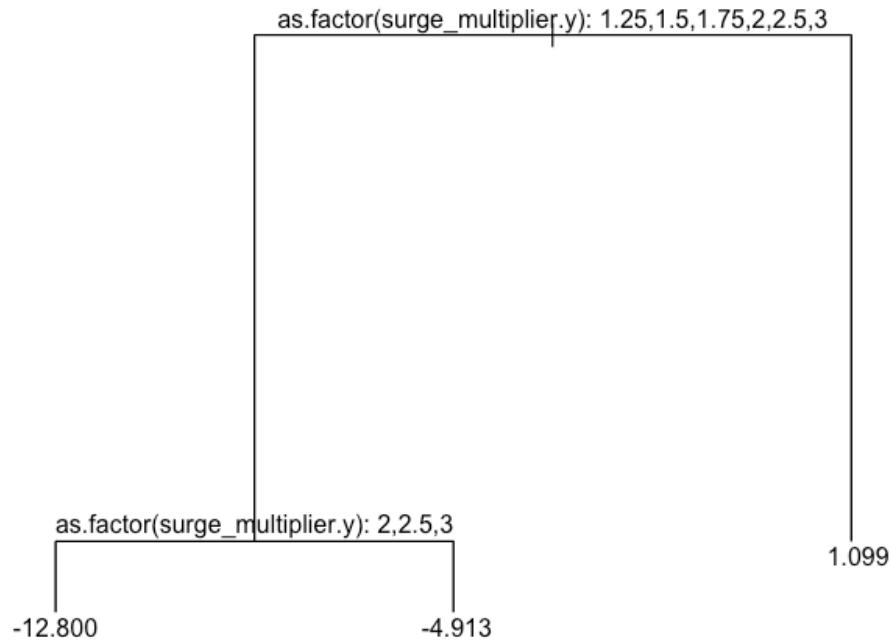


Price Difference: Uber Economy VS Lyft Economy



55.8% riding with Uber Economy is more expensive

Price Difference: Uber Economy VS Lyft Economy



MSE = 25.13 / 25.07

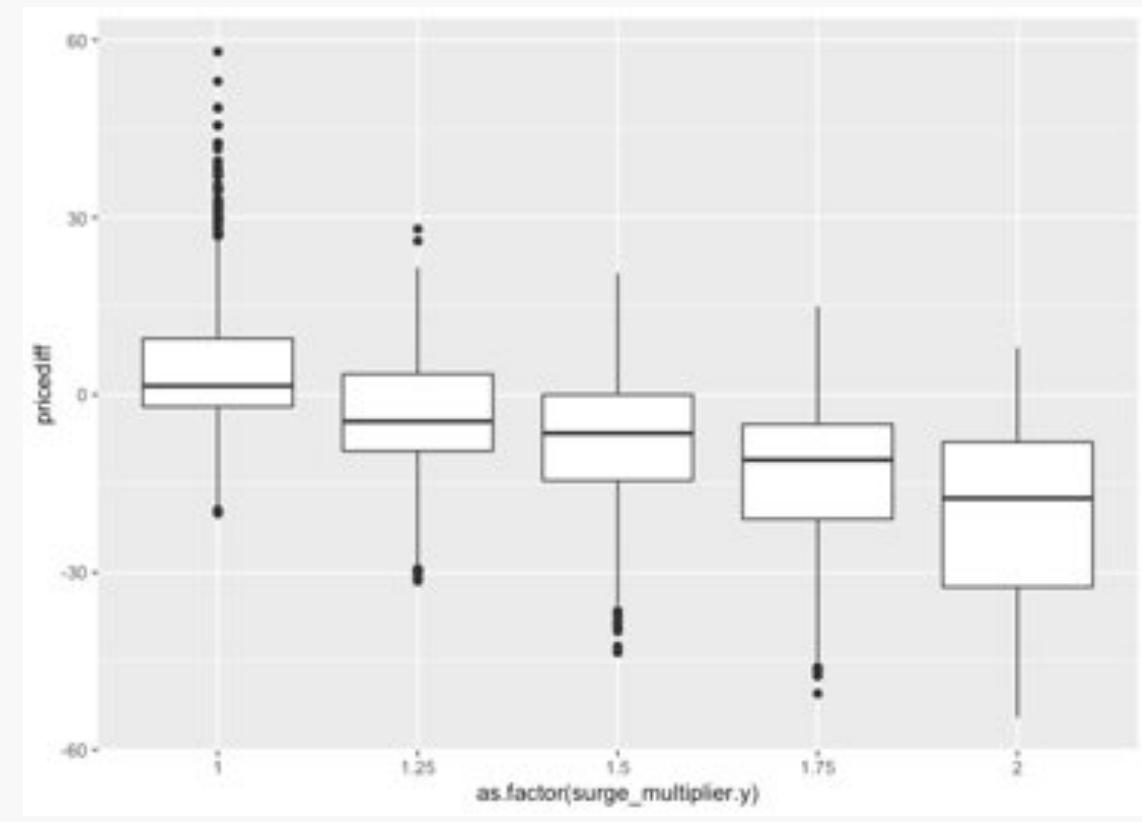
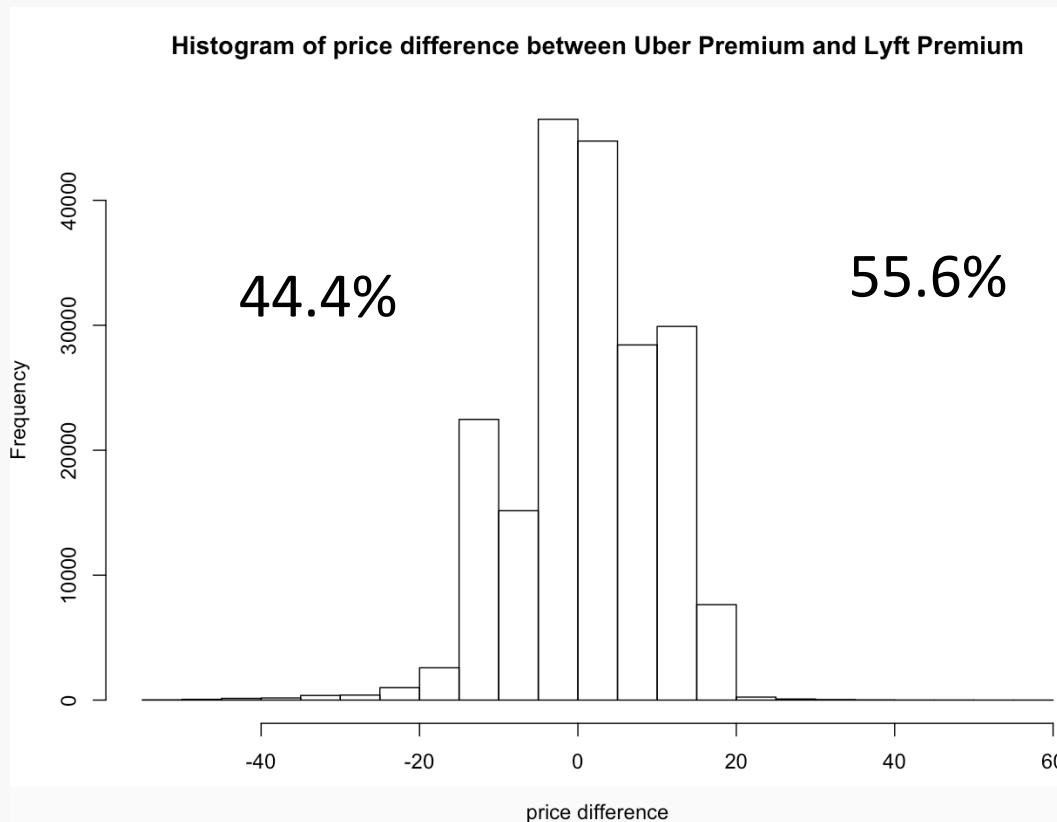
Most Important Predictor: Surge

```
Call:  
lm(formula = pricediff ~ distance + dropoff + pickup + bad_weather +  
    Isweekend + rush_hour + as.factor(surge_multiplier.y), data = samesub_e.train)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-28.964 -3.054 -0.168  2.739 60.830  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept) 1.519614  0.017981 84.512 < 2e-16 ***  
distance    -0.258942  0.009154 -28.287 < 2e-16 ***  
dropoffNearby 0.244105  0.019950 12.236 < 2e-16 ***  
pickupNearby  0.035933  0.019949  1.801 0.07166 .  
bad_weather1 -0.127838  0.033050 -3.868 0.00011 ***  
Isweekend1   0.126795  0.017601  7.204 5.86e-13 ***  
rush_hour1   -0.022586  0.040004 -0.565 0.57236  
as.factor(surge_multiplier.y)1.25 -4.835781  0.051438 -94.012 < 2e-16 ***  
as.factor(surge_multiplier.y)1.5  -6.975854  0.063351 -110.115 < 2e-16 ***  
as.factor(surge_multiplier.y)1.75 -9.090181  0.113432 -80.138 < 2e-16 ***  
as.factor(surge_multiplier.y)2   -13.215164  0.118004 -111.235 < 2e-16 ***  
as.factor(surge_multiplier.y)2.5 -12.944310  0.228995 -56.527 < 2e-16 ***  
as.factor(surge_multiplier.y)3   -28.442867  0.482556 -58.942 < 2e-16 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 4.99 on 402060 degrees of freedom  
Multiple R-squared:  0.1028, Adjusted R-squared:  0.1028  
F-statistic: 3840 on 12 and 402060 DF, p-value: < 2.2e-16
```

MSE = 24.90 / 24.86

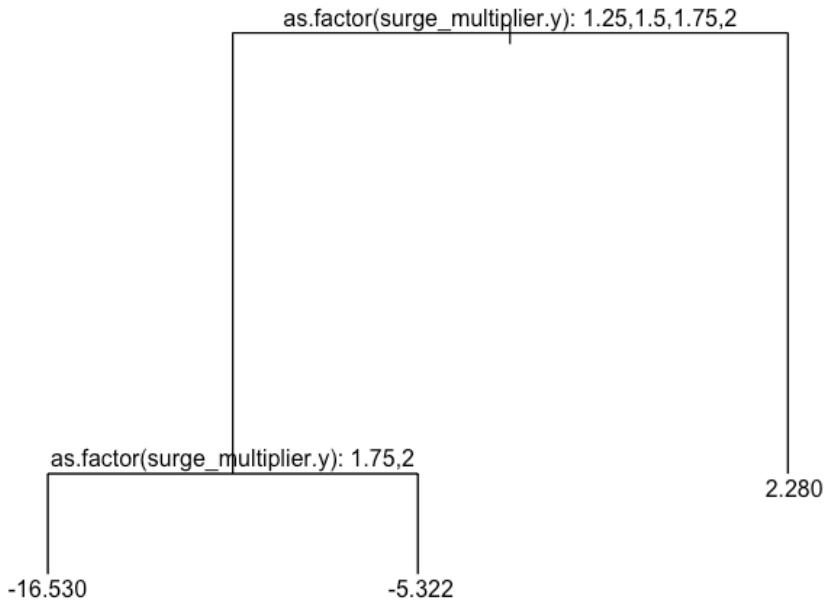
P-value = < 2.2e-16

Price Difference: Uber Premium VS Lyft Premium



Overall Uber Premium is slightly more expensive, However, there are some extreme cases when Lyft price is much higher than Uber

Price Difference: Uber Premium VS Lyft Premium



MSE = 68.93 / 68.49

Most Important Predictor: Surge

```
Call:  
lm(formula = pricediff ~ distance + dropoff + pickup + bad_weather +  
    Isweekend + rush_hour + as.factor(surge_multiplier.y), data = samesub_p.train)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-36.610  -4.414  -0.466   6.812  56.018  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)      
(Intercept)  3.13532  0.04232  74.086 < 2e-16 ***  
distance     -0.40205  0.02129 -18.885 < 2e-16 ***  
dropoffNearby  0.04033  0.04674  0.863  0.3883  
pickupNearby   -0.35631  0.04665 -7.638 2.21e-14 ***  
bad_weather1   0.16131  0.07574  2.130  0.0332 *  
Isweekend1    0.22321  0.04179  5.341 9.24e-08 ***  
rush_hour1     -0.20069  0.09132 -2.198  0.0280 *  
as.factor(surge_multiplier.y)1.25 -6.19383  0.09954 -62.228 < 2e-16 ***  
as.factor(surge_multiplier.y)1.5  -9.91040  0.12638 -78.415 < 2e-16 ***  
as.factor(surge_multiplier.y)1.75 -15.29820  0.20515 -74.571 < 2e-16 ***  
as.factor(surge_multiplier.y)2   -22.27591  0.21574 -103.255 < 2e-16 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 8.262 on 199953 degrees of freedom  
Multiple R-squared:  0.1167,    Adjusted R-squared:  0.1166  
F-statistic: 2641 on 10 and 199953 DF,  p-value: < 2.2e-16
```

MSE = 68.26 / 67.91

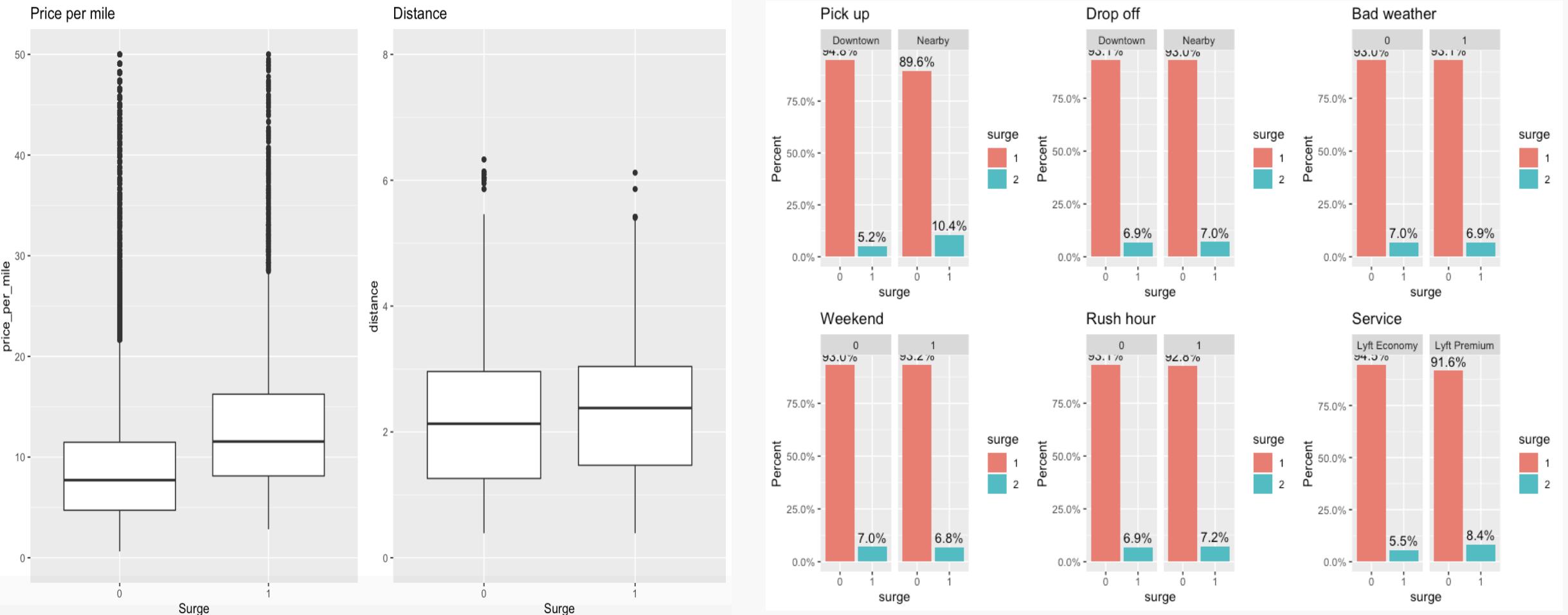
P-value = < 2.2e-16

Question 3

Is there a price per mile at which surge pricing is applied?



Useful Predictors for Logistic



- price per mile, distance, pickup ,and service would be good predictors

Logistic Regression Model - Surge

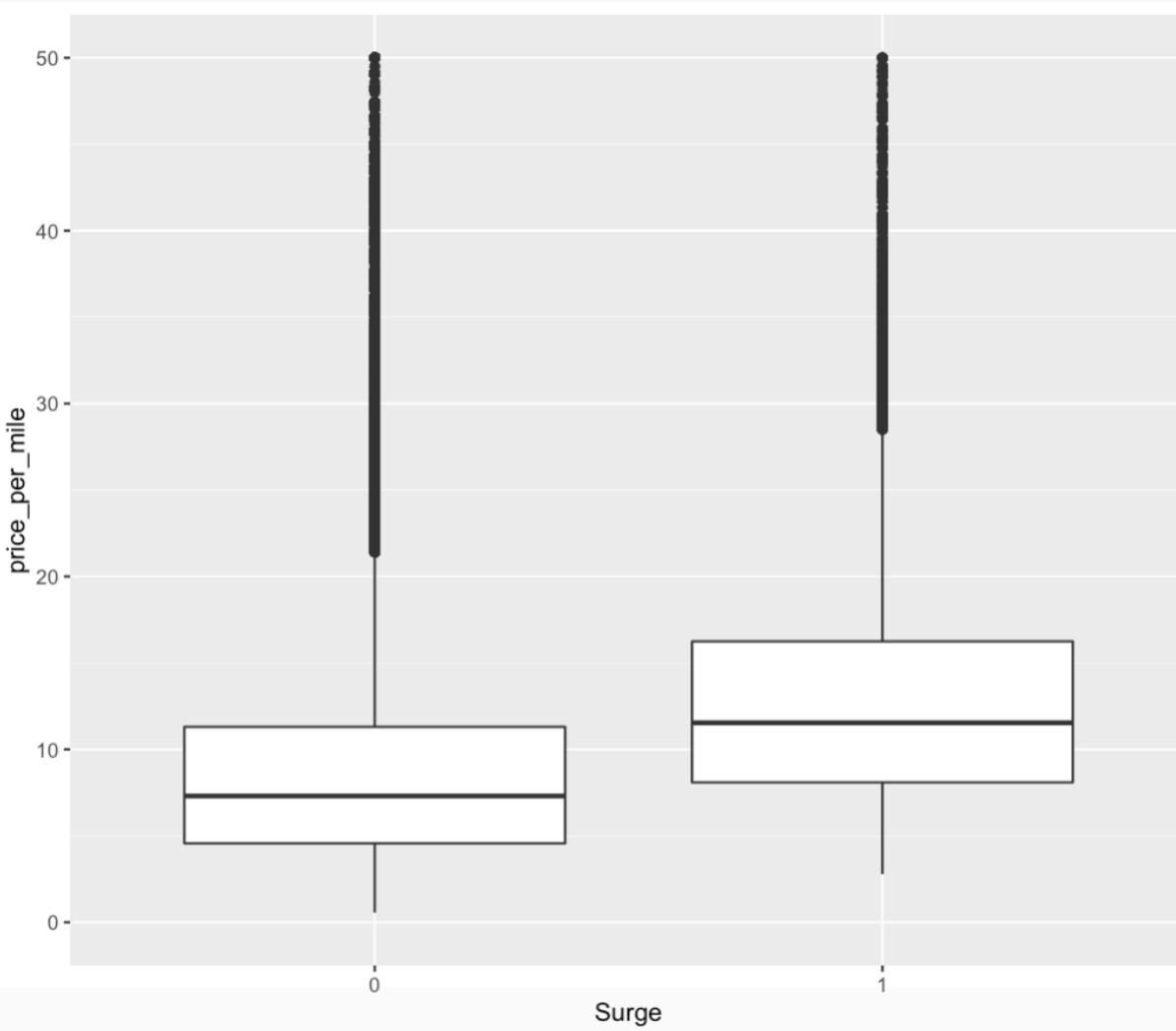
```
Call:  
glm(formula = as.factor(surge) ~ price_per_mile + service + pickup +  
  distance, family = "binomial", data = train_l)  
  
Deviance Residuals:  
    Min      1Q  Median      3Q      Max  
-1.9839 -0.4148 -0.3030 -0.2464  2.7132  
  
Coefficients:  
              Estimate Std. Error z value Pr(>|z|)  
(Intercept) -4.824178  0.038508 -125.28 <2e-16 ***  
price_per_mile  0.097485  0.001436   67.91 <2e-16 ***  
serviceLyft Premium -0.402702  0.025141  -16.02 <2e-16 ***  
pickupNearby     0.907744  0.022148   40.99 <2e-16 ***  
distance        0.436651  0.011049   39.52 <2e-16 ***  
---  
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
(Dispersion parameter for binomial family taken to be 1)  
  
Null deviance: 77232  on 153050  degrees of freedom  
Residual deviance: 70804  on 153046  degrees of freedom  
AIC: 70814  
  
Number of Fisher Scoring iterations: 6
```

- Significant factors: price per mile, service, pick up location, distance
- Confusion matrix: 93% accuracy
- Testing data includes 0.066% of surge data

glm.pred	0	1
0	142344	10102
1	458	147

```
> mean(glm.pred==test_l$surge)
[1] 0.9310034
```

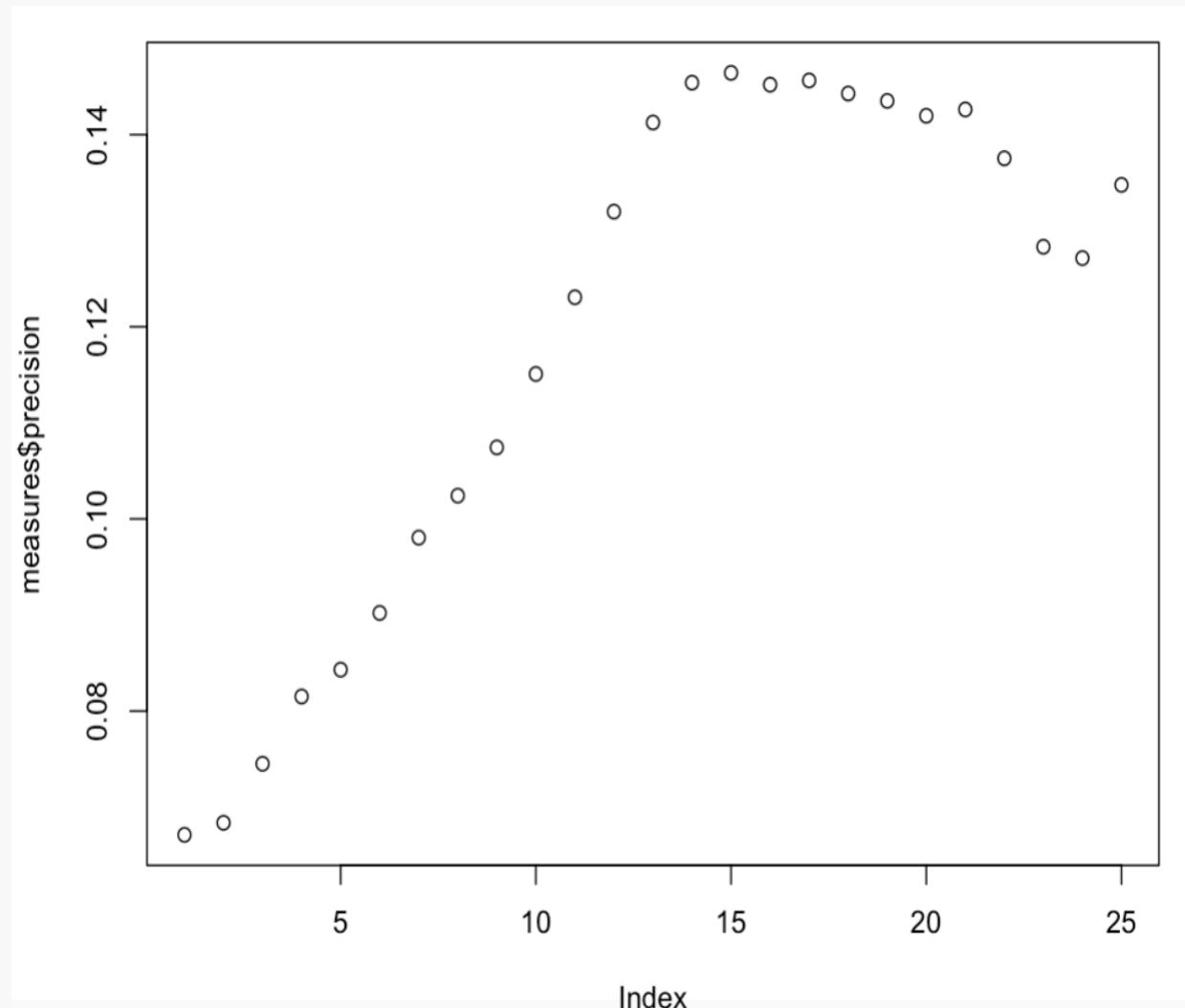
Surge VS. Price per mile



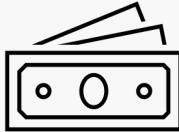
- At what level of price per mile will lead to a surge?
- Build simple model to test $\text{price per mile} > \text{\$\$\$}$: surge
 $\text{price per mile} < \text{\$\$\$}$: no surge
- Compare to logistic model

Comparison Measurement

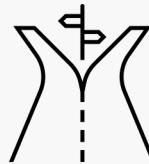
- Focus on precision rate =
$$\frac{\text{num of correct predict surge}}{\text{num of predict surge}}$$
- Best precision rate when
price per miles = 15
- Precision rate = 0.1464
- logistic precision rate = 0.24



Summary of Findings



- Best predictors for PPM in Uber is distance and car type as Uber Economy
- Best predictor for PPM in Lyft is car type as Lyft Economy and distance



- Surge was most important predictor for price difference between Uber and Lyft



- When the price per mile is over \$15, there is likely a surge in unit price
- Other factors such as pick up location, distance, and car type can also cause a surge

Limitations and Future Applications

Limitations

- Small time frame
- Didn't include surge for uber
- No extreme weather observations (small time frame)
- Missing data
- All similar area

Recommendation

- Based on our findings, someone trying to decide whether to use Uber or Lyft can observe different factors mentioned and decide which company to call for lowest price

Future Analyses

- Figure out at what price people stop calling Ubers/Lyfts and see if it is worth it to lower prices to get more customers
- Study data over longer time frame to find more bad weather observations or how holidays impact pricing
- Compare Uber and Lyft price differencing when Uber surge data is also recorded

References

<https://craft.co/lyft>

<https://ride.guru/content/newsroom/the-story-behind-uber-how-uber-started>

<https://www.ridester.com/drive-uber-lyft-boston/>

<https://www.kaggle.com/ravi72munde/uber-lyft-cab-prices>

<https://www.boston-discovery-guide.com/boston-map.html>