Most Harzardous Driving Area of Canada?

STA130 April 2018

Lucas Xian, Tony Chen, Qianshu Ni, Yang Jin

Introduction

Hundreds of thousands of Canadians get injured each year in car accidents, and thousands die each year from them. With this many accidents, it would be extremely beneficial to identify dangerous driving areas. Using the dataset "Hazardous Driving Areas" from GeoTab, we will investigate determine the most dangerous driving area in Canada, and more.

Geotab's "Hazardous Driving Areas" dataset publishes real-time and historical incident that captures both accident and near-miss events, for example, sudden braking. It provides measurements related to driving incidents, and generates a severity score to rank hazardous areas around the world.

Objectives

- The purpose of this analysis is to determine the most hazardous driving areas in Canada.
- We define an area as a hazardous driving area if its severity score is larger than the average severity score and the number of incidents is higher than the average number of incidents.

Data Summary

- **New Variables**
- Proportion of type of incidents to total incidents
- Total incidents of each type in each province
- Is an area hazardous
- **Modifications**
- Joined “Hazardous Driving Areas” dataset with a “Population of Canada” dataset

Population

```
## # A tibble: 10 x 2
##    State                          Population
##    <chr>                               <dbl>
##  1 Canada                           36963854
##  2 Newfoundland and Labrador          527613
##  3 Nova Scotia                        957470
##  4 New Brunswick                      760744
##  5 Quebec                            8439925
##  6 Ontario                          14318750
##  7 Manitoba                          1346993
##  8 Saskatchewan                      1169752
##  9 Alberta                           4318772
## 10 British Columbia                  4849442
```

Total Areas in Each Province

```
## # A tibble: 10 x 2
##    State                          total
##    <fct>                          <int>
##  1 Alberta                          477
##  2 British Columbia                 666
##  3 Manitoba                         601
##  4 New Brunswick                    120
##  5 Newfoundland and Labrador         49
##  6 Nova Scotia                      243
##  7 Ontario                         5909
##  8 Prince Edward Island               2
##  9 Quebec                          2128
## 10 Saskatchewan                      46
```

Total Number of Incidents

```
## # A tibble: 9 x 2
##   State                          total_inci
##   <fct>                               <int>
## 1 Alberta                              4968
## 2 British Columbia                     6158
## 3 Manitoba                            15412
## 4 New Brunswick                        4467
## 5 Newfoundland and Labrador             977
## 6 Nova Scotia                          6617
## 7 Ontario                            109283
## 8 Quebec                              84791
## 9 Saskatchewan                          516
```

Avgerage Severity Score and Average Number of Incidents

```
## # A tibble: 10 x 3
##    State                          avg_severity_score avg_num_
##    <fct>                                       <dbl>      <
##  1 Alberta                                     0.109      1
##  2 British Columbia                            0.0660
##  3 Manitoba                                    0.0863     2
##  4 New Brunswick                               0.130      3
##  5 Newfoundland and Labrador                   0.272      1
##  6 Nova Scotia                                 0.0603     2
##  7 Ontario                                     0.0808     1
##  8 Quebec                                      0.115      3
##  9 Saskatchewan                                0.439      1
## 10 Canada                                      0.0912     2
```

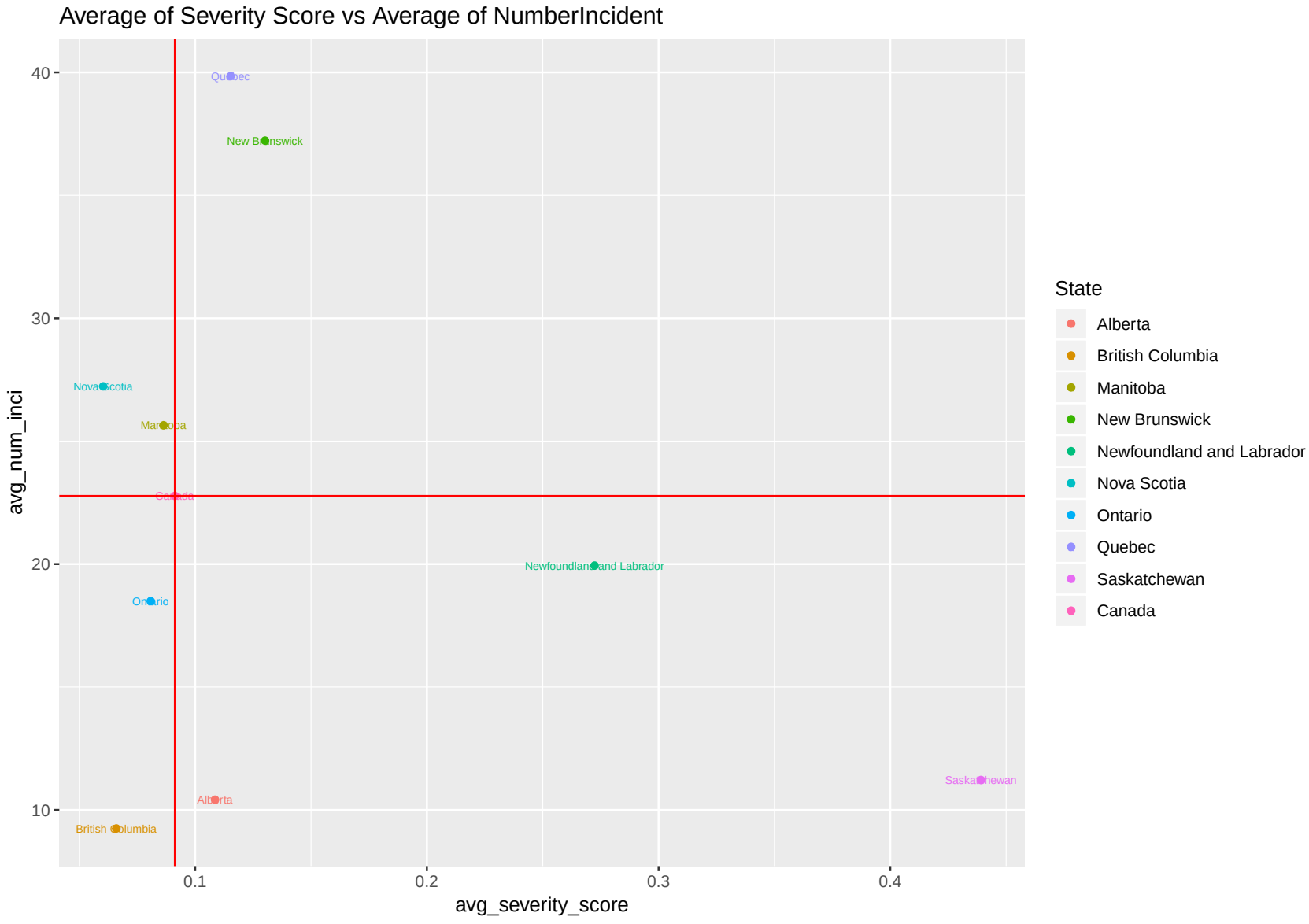
Proportion of Hazardous Area in Each Province

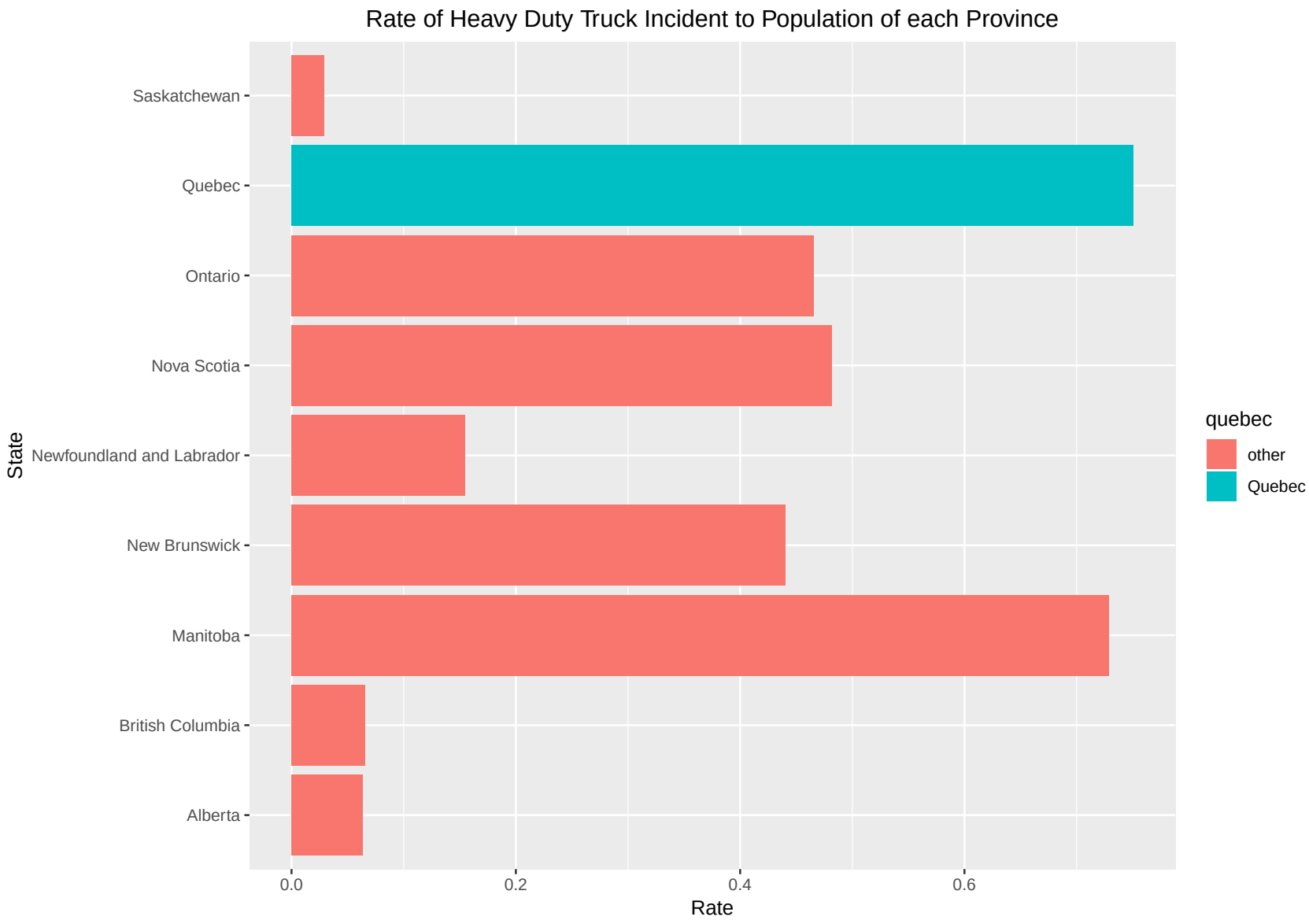
```
## # A tibble: 9 x 4
##   State                        num_yes total proportion
##   <fct>                          <int> <int>      <dbl>
## 1 Newfoundland and Labrador         28    49       57.1
## 2 Quebec                           882  2128       41.4
## 3 Saskatchewan                      19    46       41.3
## 4 New Brunswick                     46   120       38.3
## 5 Nova Scotia                       62   243       25.5
## 6 Manitoba                         153   601       25.5
## 7 Ontario                         1467  5909       24.8
## 8 Alberta                           96   477       20.1
## 9 British Columbia                 130   666       19.5
```

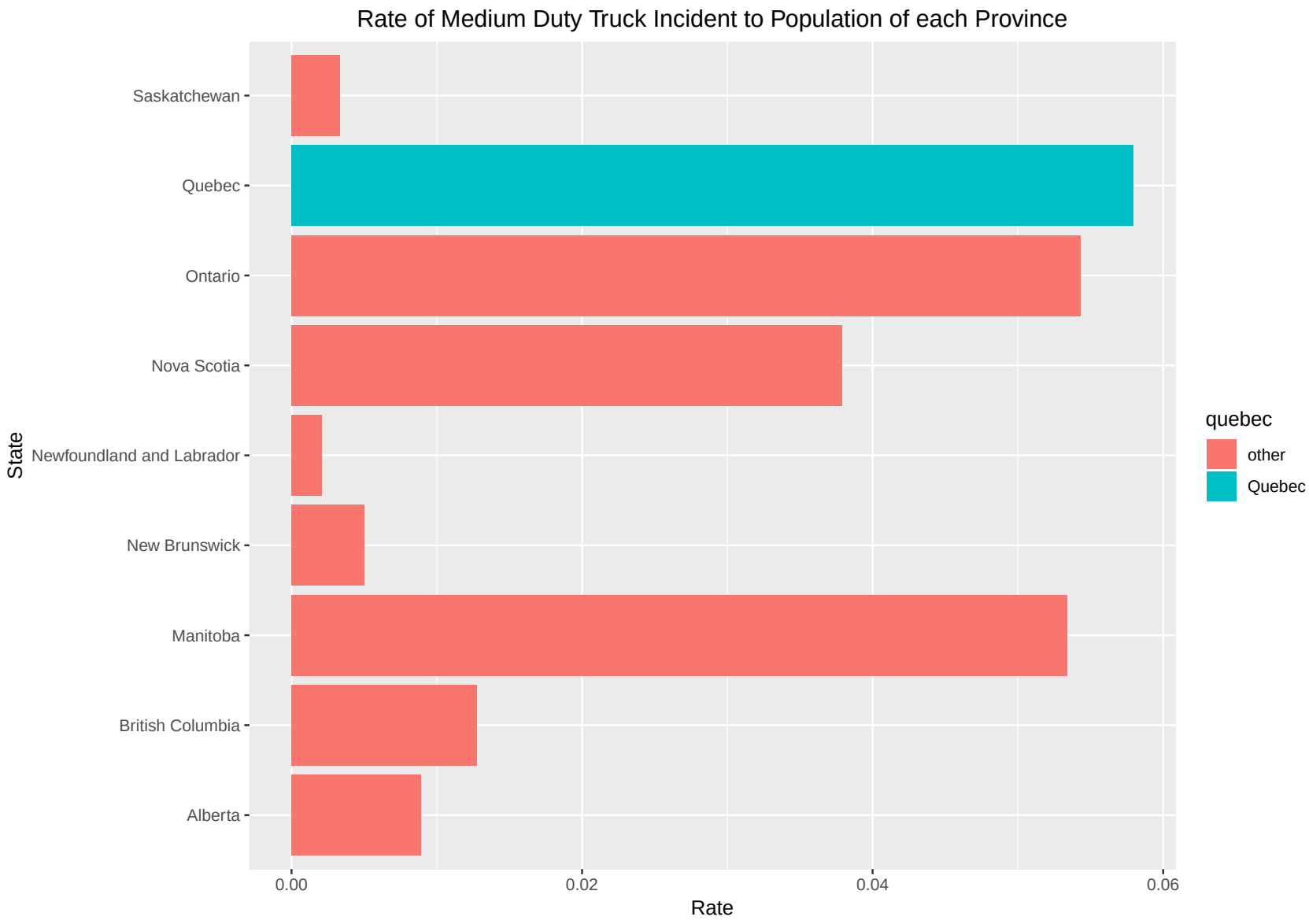
Statistical Methods

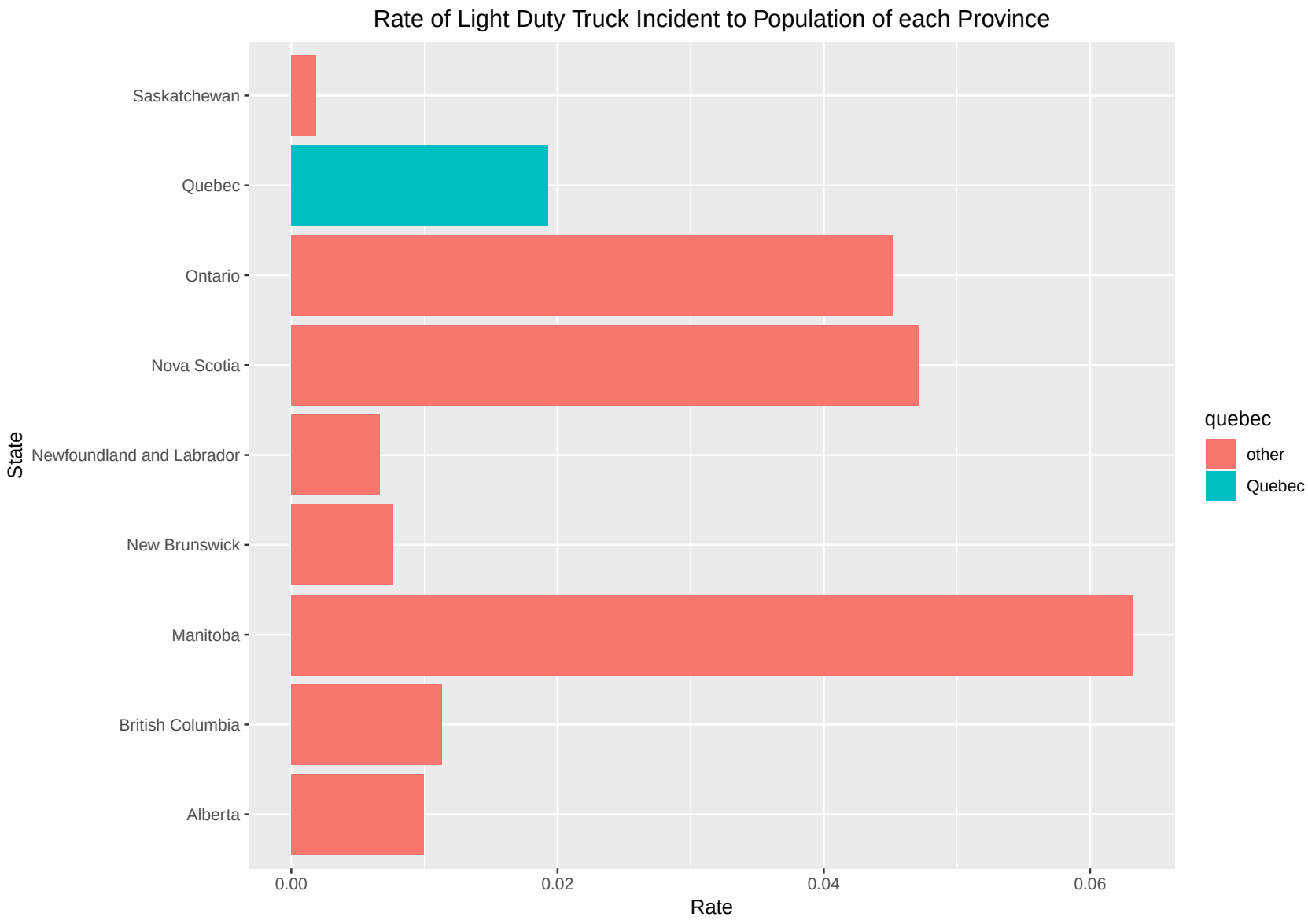
- Binary: Using ifelse to mutate a new binary variable based on our definition for hazadous area.
- ggplot: Scatterplot, Barplot
- Classification Tree: Using HdtIncident and MdtIncidentt as variables and sets three level of HdtIncident (None, Some and Many) to predict the H_driving by classification tree.
- ROC: Using the ROCR to calculate each threshold value for our classification tree

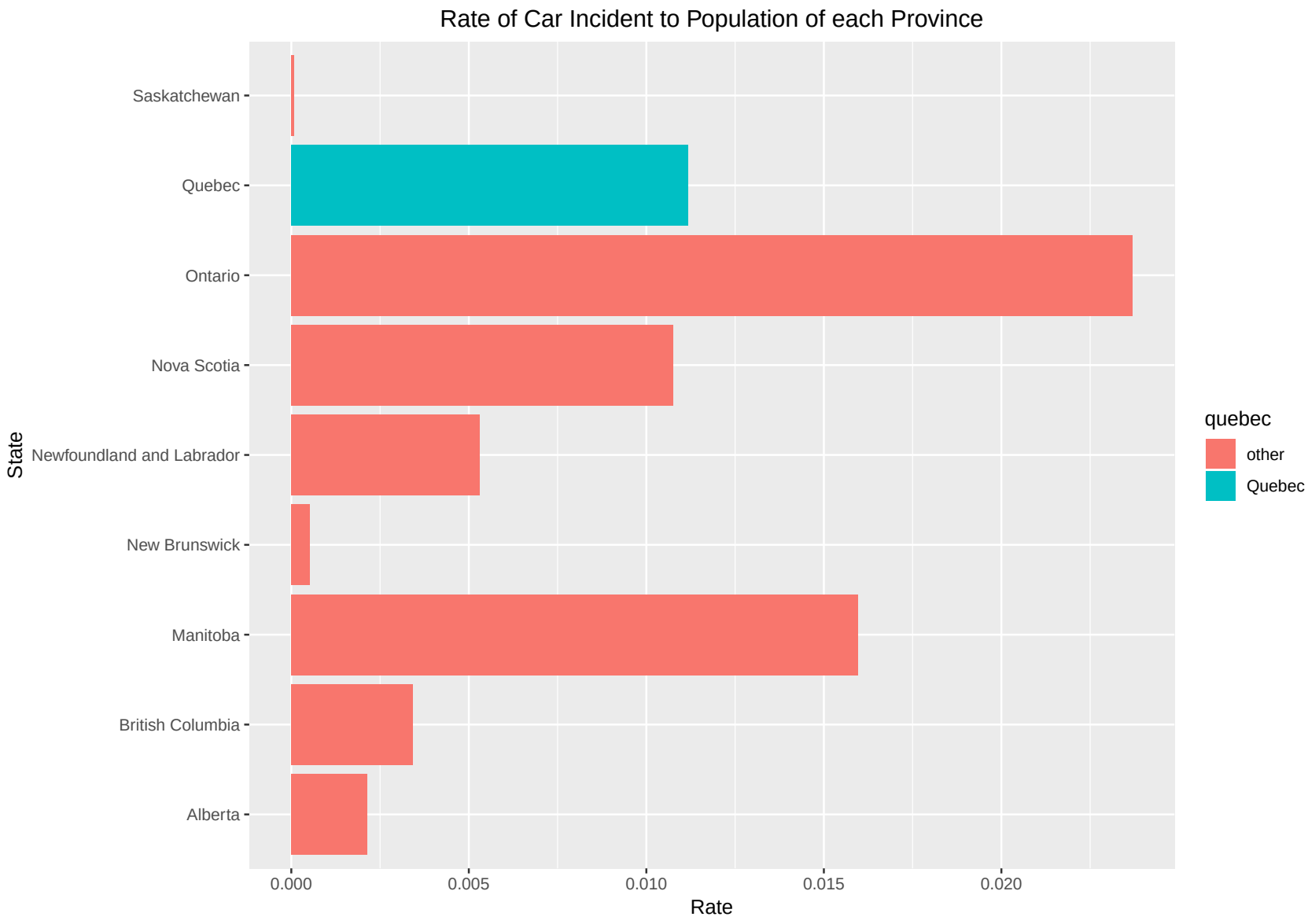
Average Sererity Score and Average Number of Incidents Scatterpolt

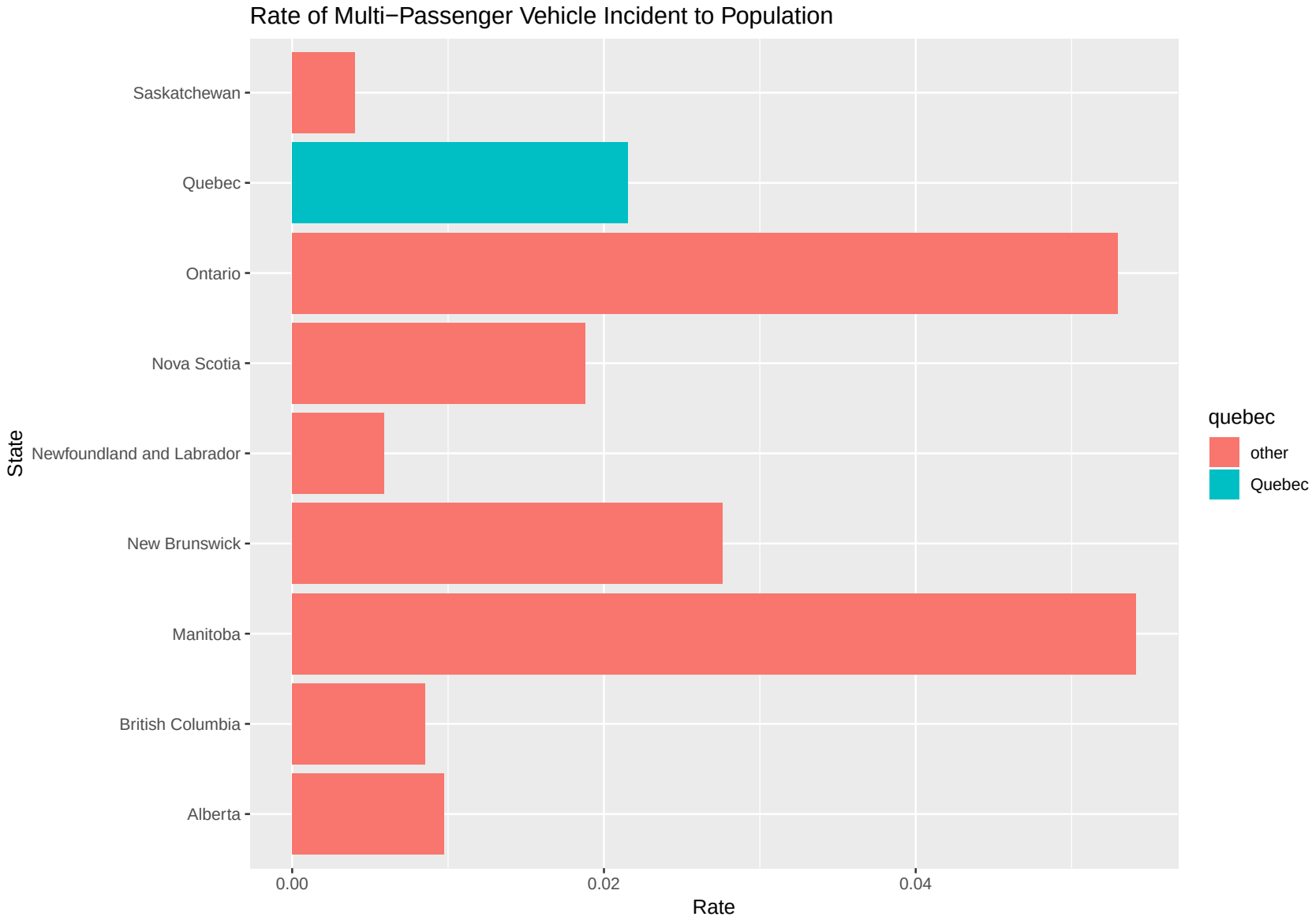



Rate of Heavy Duty Truck Incident to Population of each Province
Saskatchewan
Quebec
Ontario
Nova Scotia
Newfoundland and Labrador
New Brunswick
Manitoba
British Columbia
Alberta
State
0.0
0.2
0.4
0.6
Rate
quebec
other
Quebec

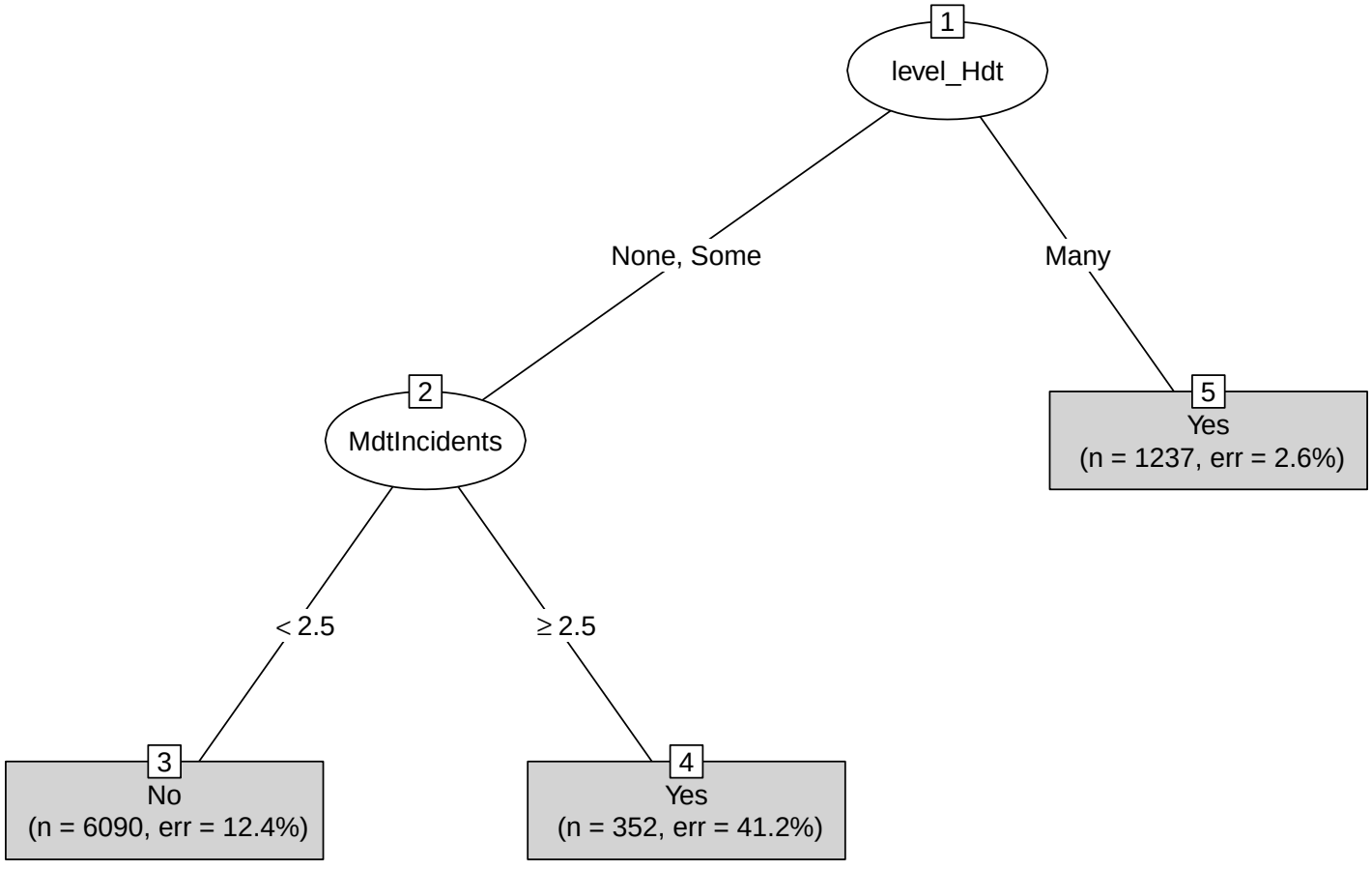

Rate of Medium Duty Truck Incident to Population of each Province
Saskatchewan
Quebec
Ontario
Nova Scotia
Newfoundland and Labrador
New Brunswick
Manitoba
British Columbia
Alberta
State
0.00
0.02
0.04
0.06
Rate
quebec
other
Quebec


Rate of Light Duty Truck Incident to Population of each Province
Saskatchewan
Quebec
Ontario
Nova Scotia
Newfoundland and Labrador
New Brunswick
Manitoba
British Columbia
Alberta
State
0.00
0.02
0.04
0.06
Rate
quebec
other
Quebec

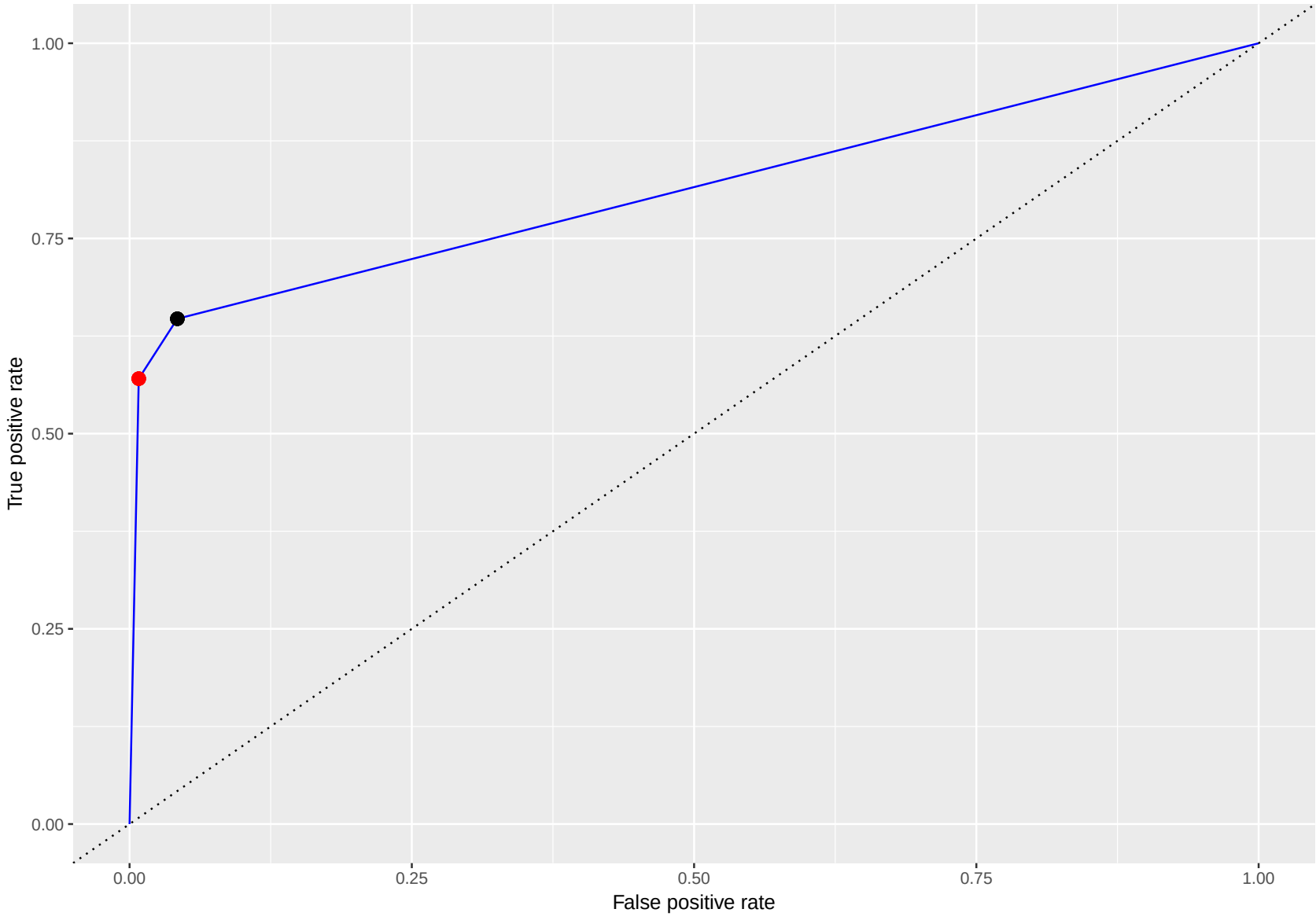

Rate of Car Incident to Population of each Province
Saskatchewan
Quebec
Ontario
Nova Scotia
Newfoundland and Labrador
New Brunswick
Manitoba
British Columbia
Alberta
State
0.000
0.005
0.010
0.015
0.020
Rate
quebec
other
Quebec


Rate of Multi−Passenger Vehicle Incident to Population
Saskatchewan
Quebec
Ontario
Nova Scotia
Newfoundland and Labrador
New Brunswick
Manitoba
British Columbia
Alberta
State
0.00
0.02
0.04
Rate
quebec
other
Quebec

Classification Tree



ROC

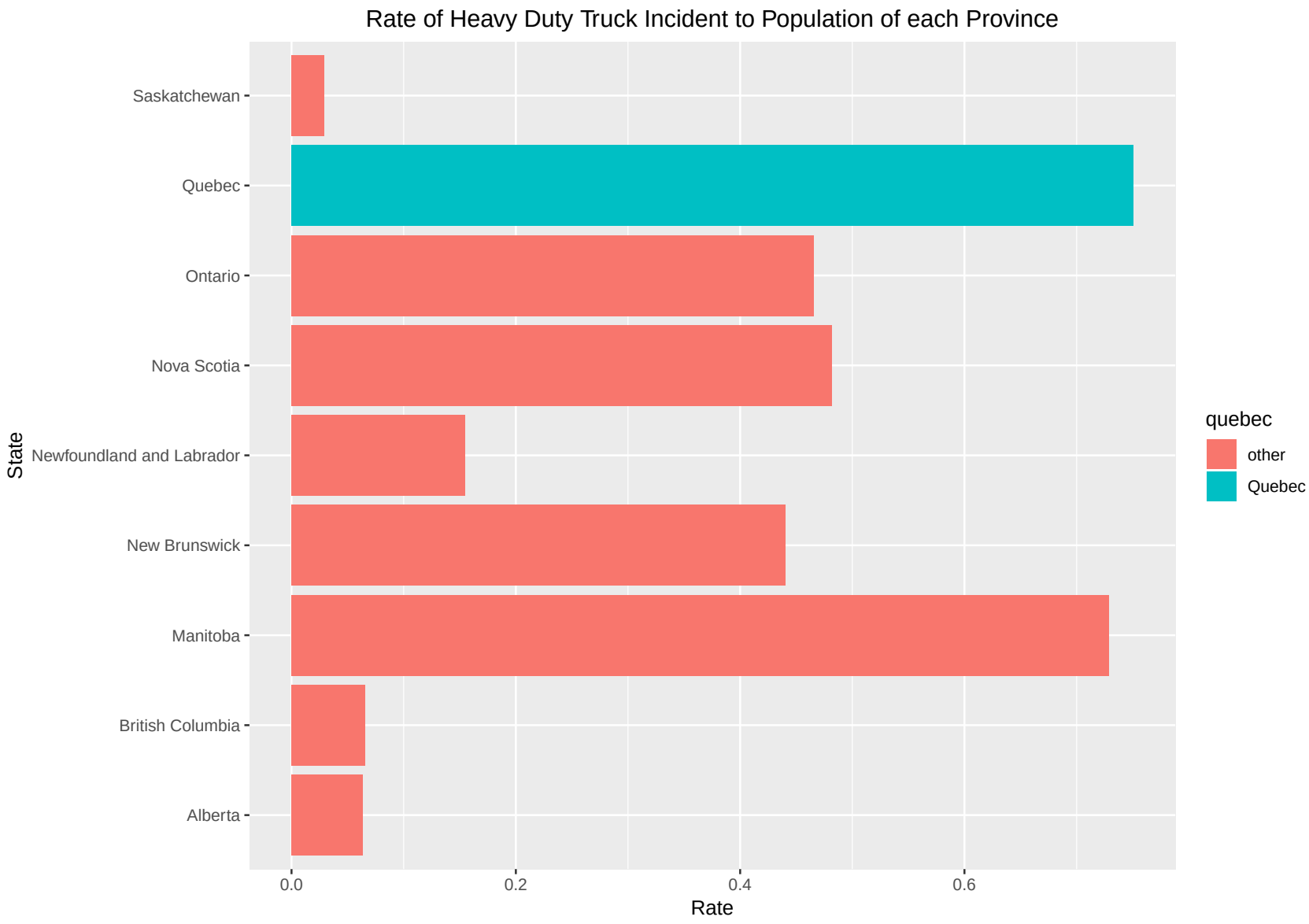

1.00
0.75
0.50
0.25
0.00
True positive rate
0.00
0.25
0.50
0.75
1.00
False positive rate

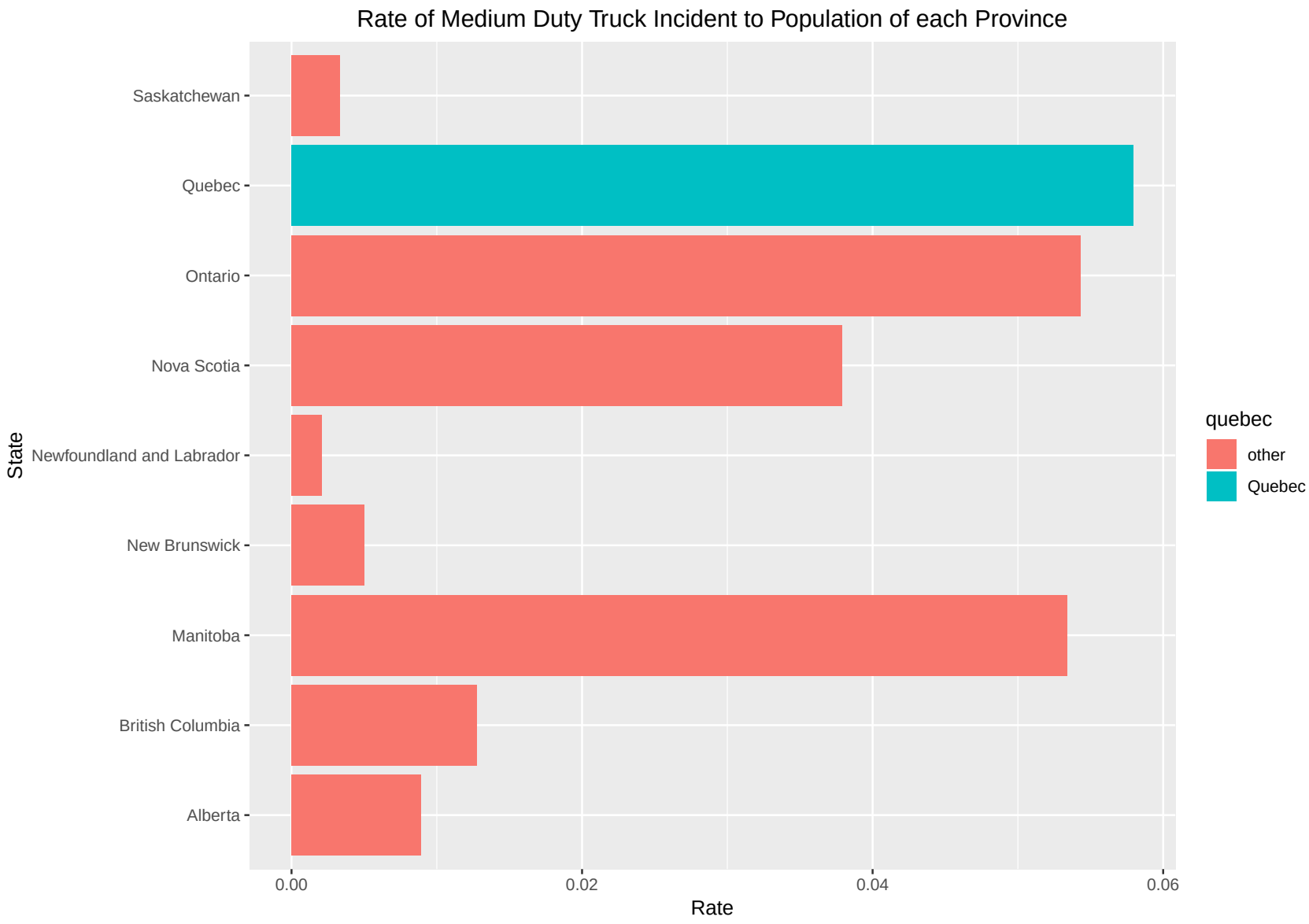
Results

- Base on the data, we discovered that the Saskatchewan and the Newfoundland and Labrador are the most two hazardous driving area
- However, since they have the smallest dataset which their datasets are less than 50, so we cannot use them to satisfy our prediction.
- As the results, we find Quebec is the most hazardous driving area.
- Classification Tree
- Two significant main effects – the number of incidents involving a medium-duty truck and a heavy-duty truck in Quebec state (from the figure P_Hdt and P_Mdt).

Proportion of Hazardous Area in Each Province

```
## # A tibble: 9 x 4
##   State                        num_yes total proportion
##   <fct>                          <int> <int>      <dbl>
## 1 Newfoundland and Labrador         28    49       57.1
## 2 Quebec                           882  2128       41.4
## 3 Saskatchewan                      19    46       41.3
## 4 New Brunswick                     46   120       38.3
## 5 Nova Scotia                       62   243       25.5
## 6 Manitoba                         153   601       25.5
## 7 Ontario                         1467  5909       24.8
## 8 Alberta                           96   477       20.1
## 9 British Columbia                 130   666       19.5
```


Rate of Heavy Duty Truck Incident to Population of each Province
Saskatchewan
Quebec
Ontario
Nova Scotia
Newfoundland and Labrador
New Brunswick
Manitoba
British Columbia
Alberta
State
0.0
0.2
0.4
0.6
Rate
quebec
other
Quebec


Rate of Medium Duty Truck Incident to Population of each Province
Saskatchewan
Quebec
Ontario
Nova Scotia
Newfoundland and Labrador
New Brunswick
Manitoba
British Columbia
Alberta
State
0.00
0.02
0.04
0.06
Rate
quebec
other
Quebec

Conclusion

- **conclusion**: Quebec is the most Hazardous driving province
- -> Highest probability of Heavy and Medium Duty Truck Incidents
- **limitation** : may only work for this specific data set
- **error** : Newfoundland and Labrador has the highest proportion of hazardous driving area, it does not match our definition of hazardous driving area

Reference

"Canadian Motor Vehicle Traffic Collision Statistics: 2015." Goverment of Canada -Transport Canada, 26 May 2017, www.tc.gc.ca/eng/motorvehiclesafety/tp-tp3322-2015-1487.html.

"Hazardous Driving Areas." Geotab Data, 13 Mar. 2018, data.geotab.com/urban-infrastructure/hazardous-driving.

Statistics Canada. "Population and Dwelling Count Highlight Tables, 2011 Census."Government of Canada, Statistics Canada, 9 Aug. 2016, www12.statcan.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=301&SR=1&S=3&O=D&RPP=25&PR=0&