



ETC513 Assignment 3: Comparison of Energy and Pollution by Country

Shaohu Chen

Master of Business Analytics(In Progress)

Qian Duan

Master of Business Analytics(In Progress)

Tina Tsou

Master of Business Analytics(In Progress)

Report for
Australian Government

**Our consultancy
add names &
add names**

📞 (03) 9905 2478
✉️ questions@company.com

ABN: 12 377 614 630

27 May 2021

Introduction

Step 1

Step 2

```
library(scales)
```

```
##  
## Attaching package: 'scales'  
  
## The following object is masked from 'package:purrr':  
##  
##     discard  
  
## The following object is masked from 'package:readr':  
##  
##     col_factor
```

The whole pass rate of Queensland

```
wholerate <- drive %>%  
  group_by(`Exam Result`) %>%  
  count() %>%  
  pivot_wider(id_cols = `Exam Result`,  
              names_from = `Exam Result`,  
              values_from = n) %>%  
  mutate(passrate = PASS/(FAIL+PASS))  
head(percent(wholerate$passrate))
```

```
## [1] "63%"
```

The driving examination pass rate of Queensland is 63% .

```
#count the pass rate of each product type

type <- drive %>%
  group_by(`Product Type Name`, `Exam Result`) %>%
  count() %>%
  pivot_wider(id_cols = `Exam Result`,
              names_from = `Exam Result`,
              values_from = n) %>%
  mutate(sum = FAIL + PASS,
        pass_rate = round(PASS /sum,2))%>%
  select(`Product Type Name`, pass_rate) %>%
  arrange(pass_rate) %>%
  mutate(pass_rate = percent(pass_rate))

knitr::kable(type, caption = "The pass rate of each product type",
             booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 1: *The pass rate of each product type*

Product Type Name	pass_rate
CLASS CA - CAR (AUTOMATIC)	53%
CLASS C - CAR (MANUAL)	56%
CLASS HR - HEAVY RIGID VEHICLE	70%
CLASS MR - MEDIUM RIGID VEHICLE	77%
CLASS RE - MOTORCYCLE (UP TO 250CC)	77%
CLASS HC - HEAVY COMBINATION VEHICLE	79%
CLASS LR - LIGHT RIGID VEHICLE	83%
CLASS R - MOTORCYCLE (OVER 250CC)	86%

```
# select the age group

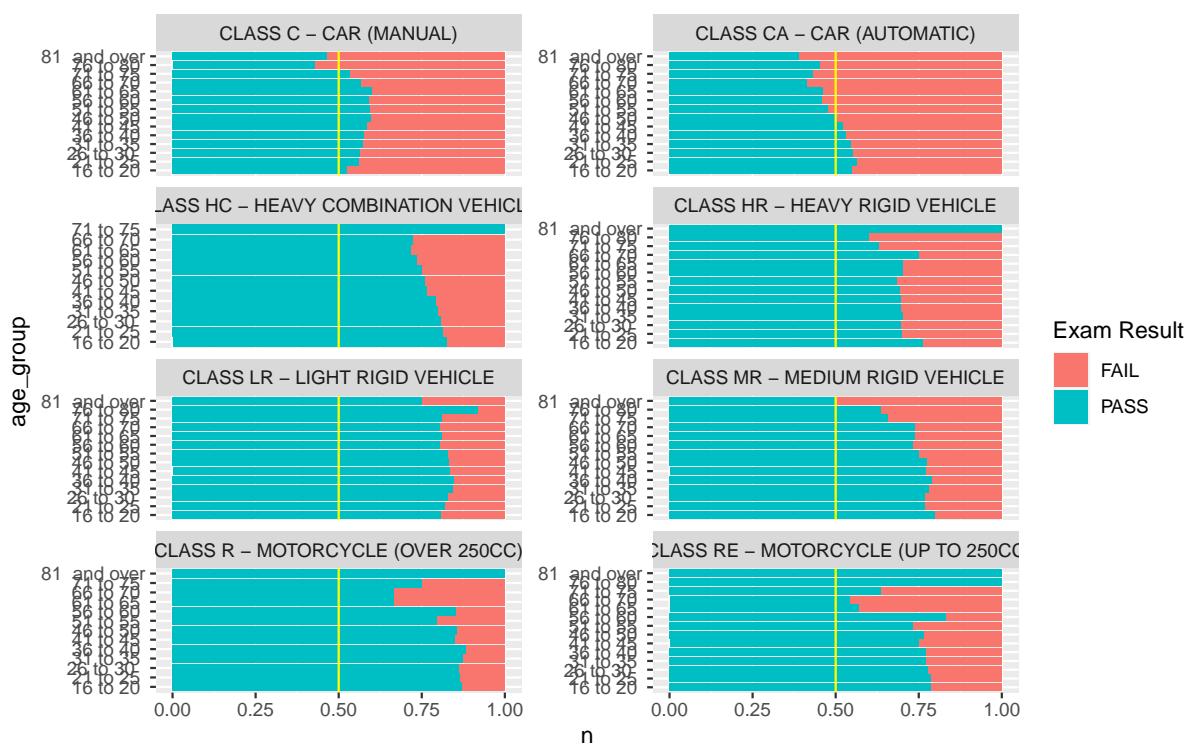
age <- drive %>%
  select(`Product Type Name`, `Driver Age Group`, `Exam Result`, `Number of Examinations`) %>%
  mutate(Age_group1 = str_remove(`Driver Age Group`, "Aged"),
        Age_group2 = str_remove(Age_group1, "years")) %>%
  rename(age_group = Age_group2) %>%
# count the pass rate by different product type
```

```
group_by(`Product Type Name`, age_group, `Exam Result`) %>%
count()
```

```
# plot the age group by license
ggplot(age)+

  geom_bar(aes(y = age_group,
                x = n,
                fill = `Exam Result`),
            stat = "identity",
            position = "fill")+

  geom_vline(xintercept = 0.5, color = "yellow")+
  facet_wrap(~`Product Type Name`, nrow = 4,
            scale = "free_y")
```



This graph illustrates all the driver license pass rate according to different age group. Most pass rate is higher than fail rate, except the manual car in older age. This is because it's too hard to pass in 80 years old. Same as automatic car with higher fail rate when the age above 45 years old.

```
# count the pass rate
pass_rate <- drive %>%
  select(`Product Type Name`, `Driver Age Group`, `Exam Result`, `Number of Examinations`) %>%
  filter(`Exam Result` == "PASS") %>%
# There are 211,717 people recorded pass
  pivot_wider(names_from = `Driver Age Group`,
              values_from = `Number of Examinations`)

## Warning: Values are not uniquely identified; output will contain list-cols.
## * Use 'values_fn = list' to suppress this warning.
## * Use 'values_fn = length' to identify where the duplicates arise
## * Use 'values_fn = {summary_fun}' to summarise duplicates
```

Step 3