

ETC 2420/5242 Lab 11 2017

Di Cook

Week 11

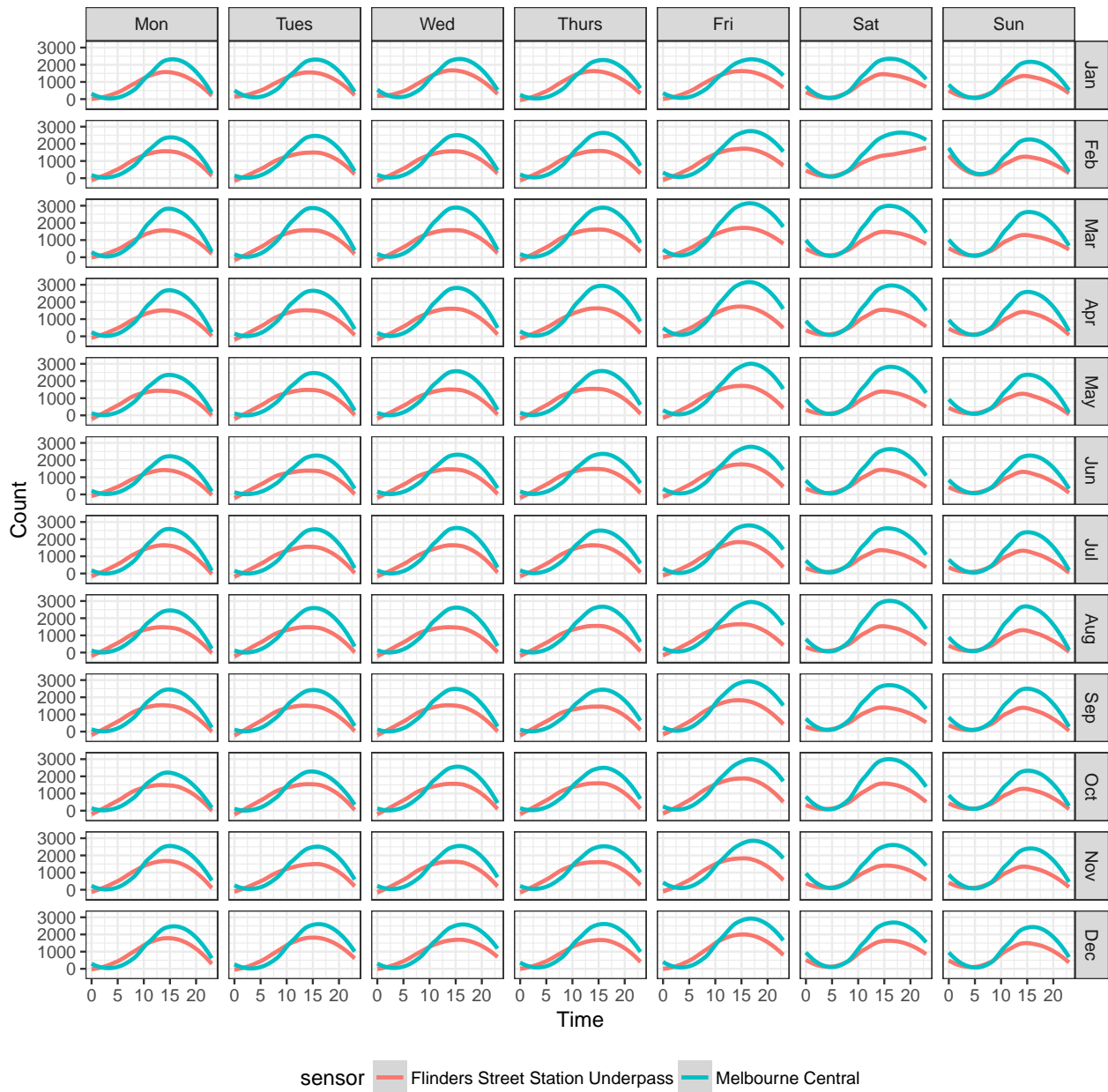
Purpose

In this lab we will model risk and loss for running a coffee shop in downtown Melbourne using the pedestrian traffic data and weather data. Data is from 2013-2014. We are going to play a game!

Models

You are provided with a fitted model for pedestrian traffic at two locations, Melbourne Central and Flinders Street Station Underpass, which was built using this code.

```
ped_weather <- read_csv("melb_ped_weather.csv")
ped_weather$time <- factor(ped_weather$time)
ped_weather$day <- factor(ped_weather$day)
ped_weather$month <- factor(ped_weather$month)
ped_weath_glm <- glm(count~day*time*month*sensor+
  high_tmp+low_tmp+high_prdp,
  data=ped_weather, family=poisson(link="log"))
save(ped_weath_glm, file="ped_weather_glm.rda")
```

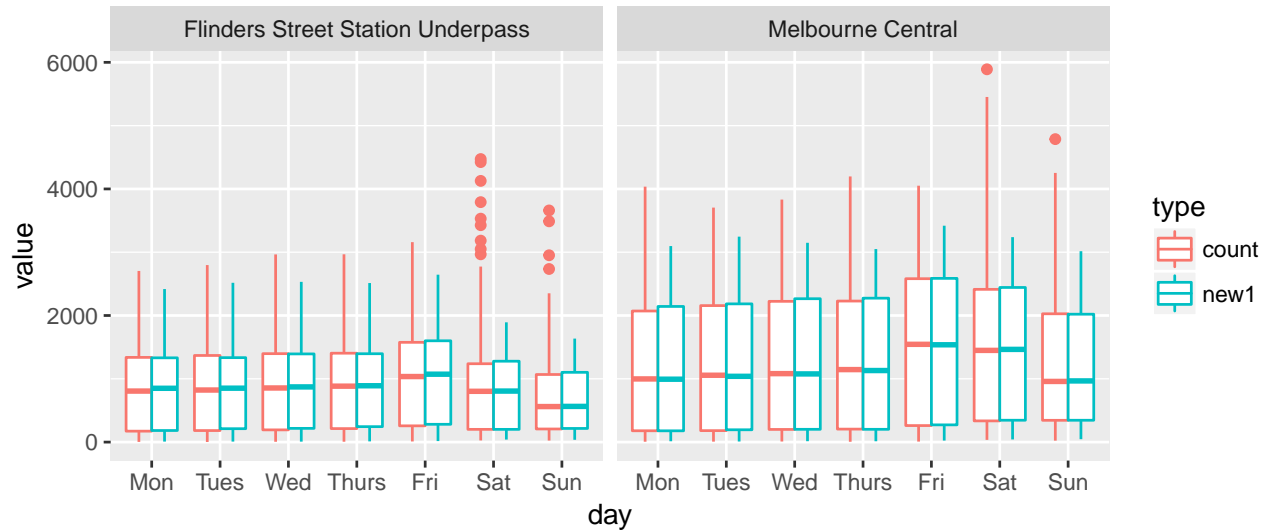


Assumptions

- At Flinders, the proportion of pedestrians passing by who will buy a coffee is 0.1 between 7-10am, 0.05 between 10-4, 0.01 between 4-8. At Melbourne Central, the proportion who will buy coffee is 0.08 between 7-10am, 0.06 between 10-4, 0.02 between 4-10pm. At all other times assume no purchases.
- Each coffee purchase is \$4.
- To open the coffee shop with one attendant costs \$100/hour. Two attendants is \$150/hour. Three attendants is \$200/hour and four attendants is \$250/hour. One attendant per hour can handle 30 customers, if the number is more than the number the staff can handle the customers will walk out without purchasing.
- You can be open no more than 8 hours per day, and the hours open can be different by day of the week. The same pattern, though, is used for every week, all months. The hours each day need to be contiguous.

Question 1

- a. (2pts) Plot the sensor traffic data, by day of the week, using a side-by-side boxplot, coloured by faceted by sensor.



- b. (4pts) Write down the objective function for the game, the difference in profit (loss) between a coffee shop at the two locations, taking into account the predicted number of pedestrians, proportion of coffee buyers, hour of day, and number of attendants.

Assume $x_{FI}(d, t)$, $x_{MC}(d, t)$ is the number of pedestrians that pass by each location for day d and time t . And $p_{FI}(t)$, $p_{MC}(t)$ is the proportion that buy a coffee at time t . Let e_{FI} , e_{MC} be the number of employees working in any hour, ranging from 0 (if closed) to 4. $y_{FI}(d, t)$, $y_{MC}(d, t)$ are the earnings for day d and time t .

$$y_{FI}(d, t) = \min(p_{FI}(t) \times x_{FI}(d, t), e_{FI} \times 50) \times 4 - 100 - (e_{FI} - 1) \times 50 \text{ if } e_{FI} > 0$$

Equivalently for $y_{MC}(d, t)$.

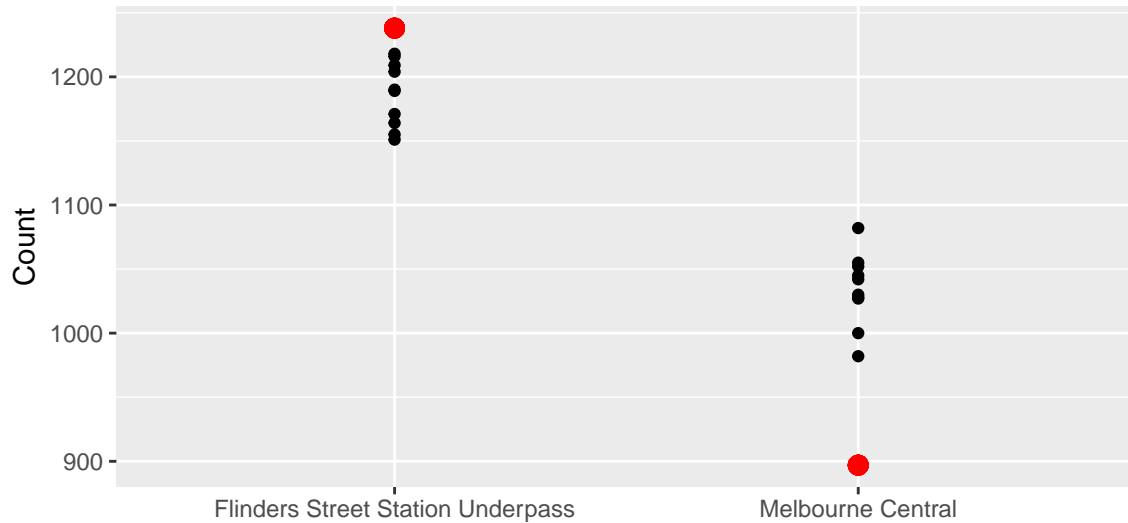
- c. (2pts) Predict the earnings for both locations for the day “2014-03-17” 8am, for all combinations of 1-4 employees. What is the optimal number of employees at each location? 4 for FI, 2 for MC

Question 2

- a. Simulate 10 sets of new pedestrian counts using the `simulate` function.
 b. Subset to examine only records for March 28, 2013, 9am. (What day of the week is this?)

```
# [1] Thurs
# Levels: Sun < Mon < Tues < Wed < Thurs < Fri < Sat
```

- c. (2pts) Summarise the distribution of the 10 values for the two locations, and compare with the actual count. Flinders has higher count. Actual counts are outside the range of the simulated values.



d. (2pts) If you are open, have three attendants at Flinders and two attendants at Melbourne Central, how much would you make at most, and at least under these conditions, at each location?

```
# # A tibble: 2 x 3
#           sensor `min(value)` `max(value)`
#           <chr>     <dbl>     <dbl>
# 1 Flinders Street Station Underpass      1151      1218
# 2 Melbourne Central                982      1082
```

Flinders it would be $1218 \cdot 1 \cdot 4 - 200 = 287$ at highest and $1151 \cdot 1 \cdot 4 - 200 = 260$ at the lowest. Melbourne Central

Question 3

a. (2pts) Now extend this to the full day (between 7am-10pm, closing at 10pm), keep the same number of attendants for the full day. How much in profit do you make at most, and at the lowest at each location?

```
# [1] 236 400 252 -28 4 152 156 64 40 -136 -108 -140 -160 0
# [15] 0
# [1] -6 186 190 146 250 250 250 250 250 62 78 66 38 30 -18
# [1] 732
# [1] 2022
```

b. (2pts) Suppose that the weather for the day is actually a hot day. How does this affect your profits? (Predict the counts for both locations for the hot and not hot day, and compute the difference. Subtract this number from your simulation values - because these were for a not hot day.)

We need to modify the `compute_earnings` function to take the reduction in pedestrians based on the weather into account.

Question 4

a. (3pts) Now scale your calculations up for the full month of March (assuming that you are open 7 days a week). How much do you earn at most, and at worst? At both locations, assuming the same weather conditions as in the given data.

```
# [1] 10516
# [1] 57872
```

- b. (2pts) During the month your coffee machine breaks, and you need to buy a new one. The new one costs \$20000. Can you afford it?

Yes, at Melbourne Central, with the month's earnings, but not at Flinders St which didn't make enough in the month.

Question 5

(3pts for reporting results) Let's play the game.

- Two players (two teams compete against each other)
- Records for two locations in Melbourne are used: "Melbourne Central", "Flinders Street Station Underpass". Pick which location you want to open a coffee shop. If the teams cannot decide, flip a coin.
- Decide what hours you will open your coffee shop, and how many attendants you will have working for each hour. You can have different numbers of attendees for each hour. You can be open no more than 8 hours per day, and the hours open can be different by day of the week. The same pattern, though, is used for every week, all months. (Mark your open times in the table below, to keep each team honest.)
- You will play 10 rounds of the game, which will randomly select a day, time of day and simulate the pedestrian counts at each site. You will need to tally up costs and profits, using the function `compute_earnings` provided. (Fractions are rounded to the nearest integer.)
- Winning team is the one with more earnings.

	Flinders St						
	Mon	Tue	Wed	Thu	Fri	Sat	Sun
7							
8							
9							
10							
11							
12							
1							
2							
3							
4							
5							
6							
7							
8							
9							

	Melbourne Central						
	Mon	Tue	Wed	Thu	Fri	Sat	Sun
7							
8							
9							
10							
11							
12							
1							
2							
3							
4							
5							
6							
7							
8							
9							

TURN IN

- Your `.Rmd` file
- Your `html` file that results from knitting the `Rmd`.

Game code

```
sample_day_time <- function() {
  sel_date <- sample(ped_weather$date, 1)
  sel_time <- sample(7:21, 1)
  return(list(sel_date, sel_time))
}
```

```

}
dt <- sample_day_time()
dt
compute_earnings <- function(ped_sim, sel_date, sel_time, Fl_attendants=1, MC_attendants=1) {
  Fl_rate <- c(rep(0.1, 3), rep(0.05, 6), rep(0.01, 4), rep(0, 2))[sel_time-6]
  MC_rate <- c(rep(0.08, 3), rep(0.06, 6), rep(0.02, 6))[sel_time-6]
  ped_sim <- ped_sim %>% dplyr::filter(date == sel_date, time == sel_time)
  Fl_count <- ped_sim %>%
    dplyr::filter(sensor == "Flinders Street Station Underpass") %>%
    select(new1)
  MC_count <- ped_sim %>%
    dplyr::filter(sensor == "Melbourne Central") %>%
    select(new1)
  Fl <- 0
  MC <- 0
  if (nrow(Fl_count) > 0) Fl <- min(round(Fl_count*Fl_rate, 0), 50*Fl_attendants)
  if (nrow(MC_count) > 0) MC <- min(round(MC_count*MC_rate, 0), 50*MC_attendants)
  Fl_earn <- 0
  MC_earn <- 0
  if (Fl > 0) Fl_earn <- Fl*4 - (100+50*(Fl_attendants-1))
  if (MC > 0) MC_earn <- MC*4 - (100+50*(MC_attendants-1))
  return(list(Fl_earn, MC_earn))
}
ped_sim <- ped_weather
ped_sim$new1 <- simulate(ped_weath_glm)$sim_1
compute_earnings(ped_sim, dt[[1]], dt[[2]], 1, 2)

```