

```
1 library(dplyr)
2 library(lubridate)
3 library(readr)
4 library(ggplot2)
5 library(tidyr)
6 library(zoo)
7
8 gett <- read_csv("Gett_Trips_NYC_2016.csv") %>%
9   mutate(
10     week_start = mdy(`Pickup Start Date`),
11     month = floor_date(week_start, "month")
12   )
13
14 gett_weekly <- gett %>%
15   group_by(
16     week_start
17   ) %>%
18   summarise(
19     gett = sum(`Total Dispatched Trips`)
20   )
21
22 gett_monthly <- gett %>%
23   group_by(
24     month
25   ) %>%
26   summarise(
27     gett = sum(`Total Dispatched Trips`)
28   )
29
30 lyft <- read_csv("Lyft_Trips_NYC_2016.csv") %>%
31   mutate(
32     week_start = mdy(`Pickup Start Date`),
33     month = floor_date(week_start, "month")
34   )
35
36 lyft_weekly <- lyft %>%
37   group_by(
38     week_start
39   ) %>%
40   summarise(
41     lyft = sum(`Total Dispatched Trips`)
42   )
43
44 lyft_monthly <- lyft %>%
45   group_by(
46     month
47   ) %>%
48   summarise(
49     lyft = sum(`Total Dispatched Trips`)
50   )
51
52 uber <- read_csv("Uber_Trips_NYC_2016.csv") %>%
53   mutate(
54     week_start = mdy(`Pickup Start Date`),
55     month = floor_date(week_start, "month")
56   )
```

```
57
58 uber_weekly <- uber %>%
59   group_by(
60     week_start
61   ) %>%
62   summarise(
63     uber = sum(`Total Dispatched Trips`)
64   )
65
66 uber_monthly <- uber %>%
67   group_by(
68     month
69   ) %>%
70   summarise(
71     uber = sum(`Total Dispatched Trips`)
72   )
73
74 yellow_monthly <- read_csv("data_reports_monthly_indicators_yellow (3).csv") %>%
75   mutate(
76     month = ymd(paste0(`Month/Year`, "-01")),
77     yellow = `Trips Per Day` * days_in_month(month)
78   ) %>%
79   select(
80     month,
81     yellow
82   )
83
84 merged_weekly <- uber_weekly %>%
85   full_join(lyft_weekly, by = c("week_start" = "week_start")) %>%
86   full_join(gett_weekly, by = c("week_start" = "week_start")) %>%
87   arrange(week_start)
88
89 merged_monthly <- yellow_monthly %>%
90   full_join(uber_monthly, by = c("month" = "month")) %>%
91   full_join(lyft_monthly, by = c("month" = "month")) %>%
92   full_join(gett_monthly, by = c("month" = "month")) %>%
93   arrange(month)
94
95 merged_weekly %>%
96   filter(week_start >= ymd("2015-01-01")) %>%
97   mutate(
98     uber = uber / 1000000,
99     lyft = lyft / 1000000,
100    gett = gett / 1000000
101  ) %>%
102   write_csv('merged_weekly.csv', na = "null")
103
104 merged_monthly %>%
105   filter(month >= ymd("2013-01-01")) %>%
106   mutate(
107     yellow = yellow / 1000000,
108     uber = uber / 1000000,
109     lyft = lyft / 1000000,
110     gett = gett / 1000000
111  ) %>%
112   write_csv('merged_monthly.csv', na = "null")
113
```

```
114 merged_weekly_flat <- merged_weekly %>%  
115   gather(provider, total_trips, uber:gett)  
116  
117 merged_monthly_flat <- merged_monthly %>%  
118   gather(provider, total_trips, yellow:gett)  
119  
120 ggplot(merged_weekly_flat) +  
121   geom_line(aes(x = week_start, y = total_trips, color = provider)) +  
122   scale_y_continuous(labels = scales::comma)  
123  
124 ggplot(merged_monthly_flat) +  
125   geom_line(aes(x = month, y = total_trips, color = provider)) +  
126   scale_y_continuous(labels = scales::comma)  
127
```