# ULTIMATE CHALLENGE PART 1 REPORT

SCOTT ATKINSON

We examine the login patterns for our app in a fixed geographical location over the course of approximately three months. To do this, we group the raw timeseries into fifteen-minute intervals. At first, we do a preliminary check of the quality of the data. The box plot for the distribution of the number of logins in Figure 1 indicates quite a few outliers. Nonetheless, all the large values are plausible.

We next turn to look at an "average day" by averaging the number of logins for each time period over every day. Figure 2 shows overall behavior. There are peaks on both sides of midnight and around noon. There is a deep valley between 6AM and 10AM and another not-as-deep valley around 6PM.

We explore this pattern further by asking if the behavior can differ according to the day of the week. Let us consider the bar chart in Figure 3 of the average number of total logins simply grouped by the day of the week. Apparently, the day of the week may have some influence on login behavior. So we look at an "average week" in Figure 4. It is more apparent now that the patterns of the demand found in the data can be subdivided into a weekday pattern and a weekend pattern. We condense the above plots into the single plot in Figure 5. To summarize, on the weekdays logins peak around noon and midnight; on weekends logins reach similar levels at noon and midnight, and there is an additional surge of behavior going through midnight into the early morning hours.
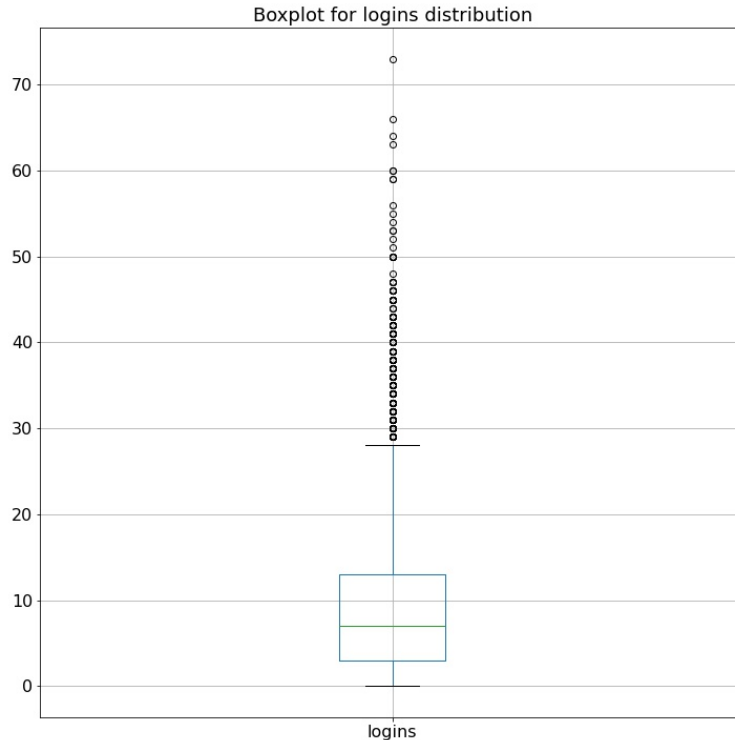


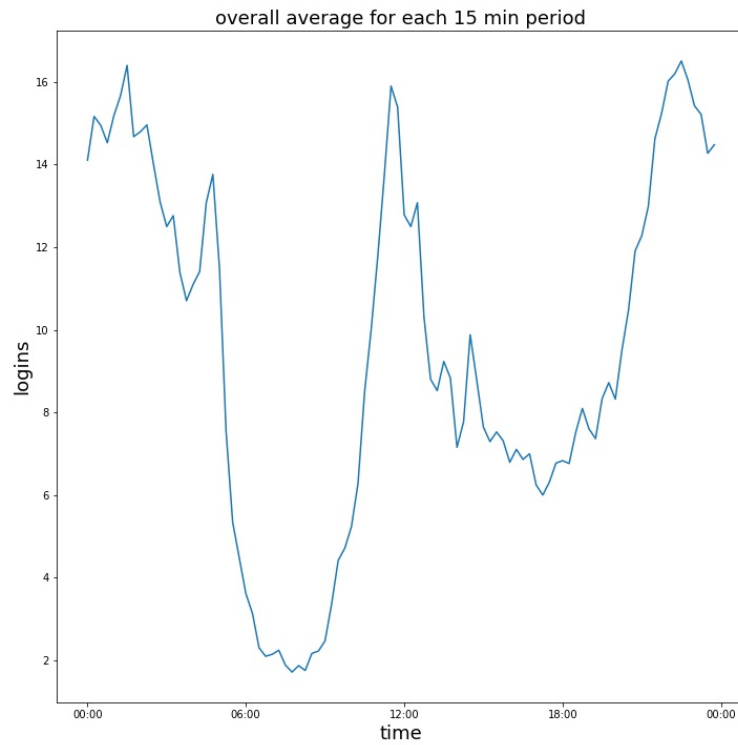FIGURE 1. Box plot of the distribution for logins

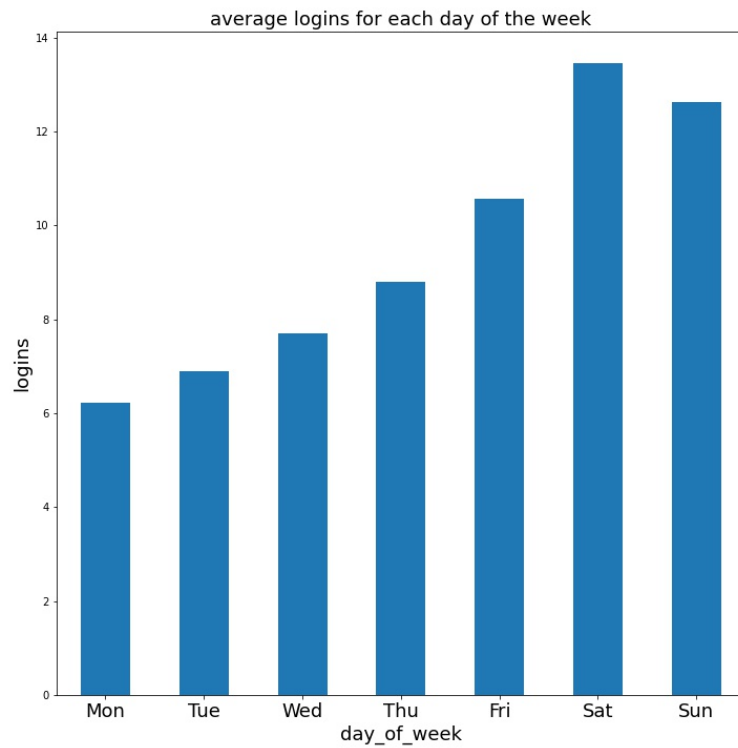FIGURE 2. Plot of the average number of logins per day for each fifteen-minute time interval



FIGURE 3. Average number of total logins grouped by the day of the week
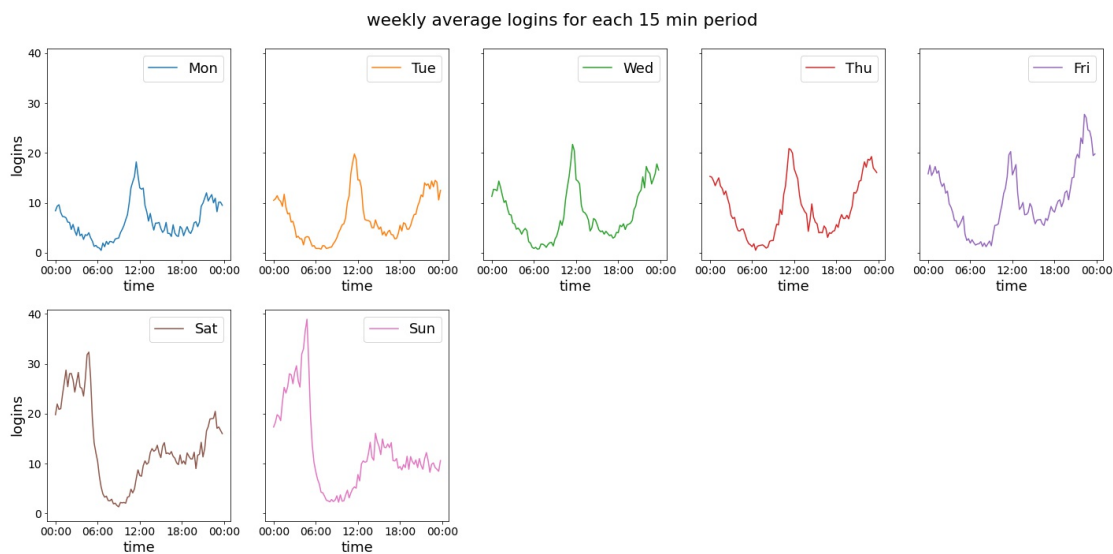
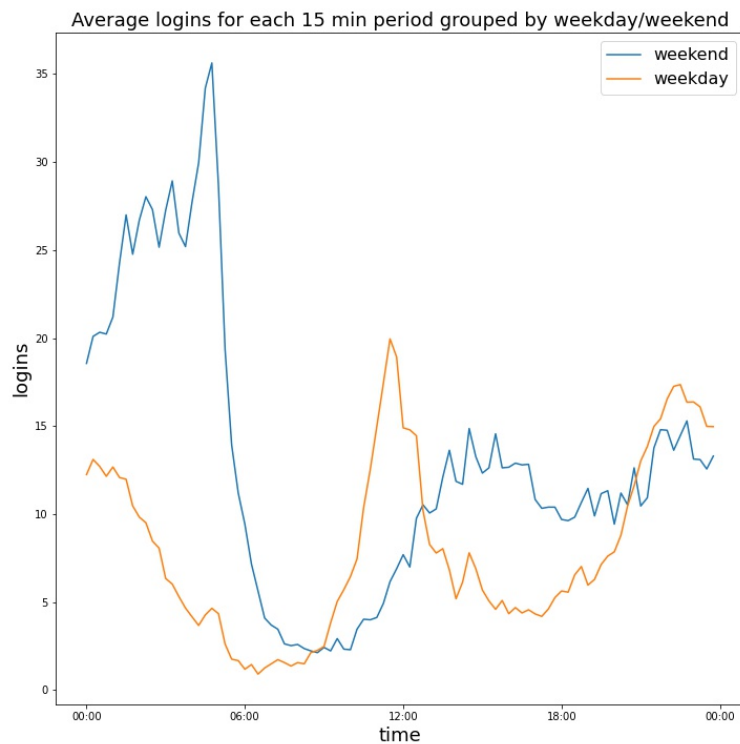FIGURE 4. Average logins for each fifteen-minute period across the week



FIGURE 5. Comparing the average logins for each 15-minute period on a weekday vs. on the weekend