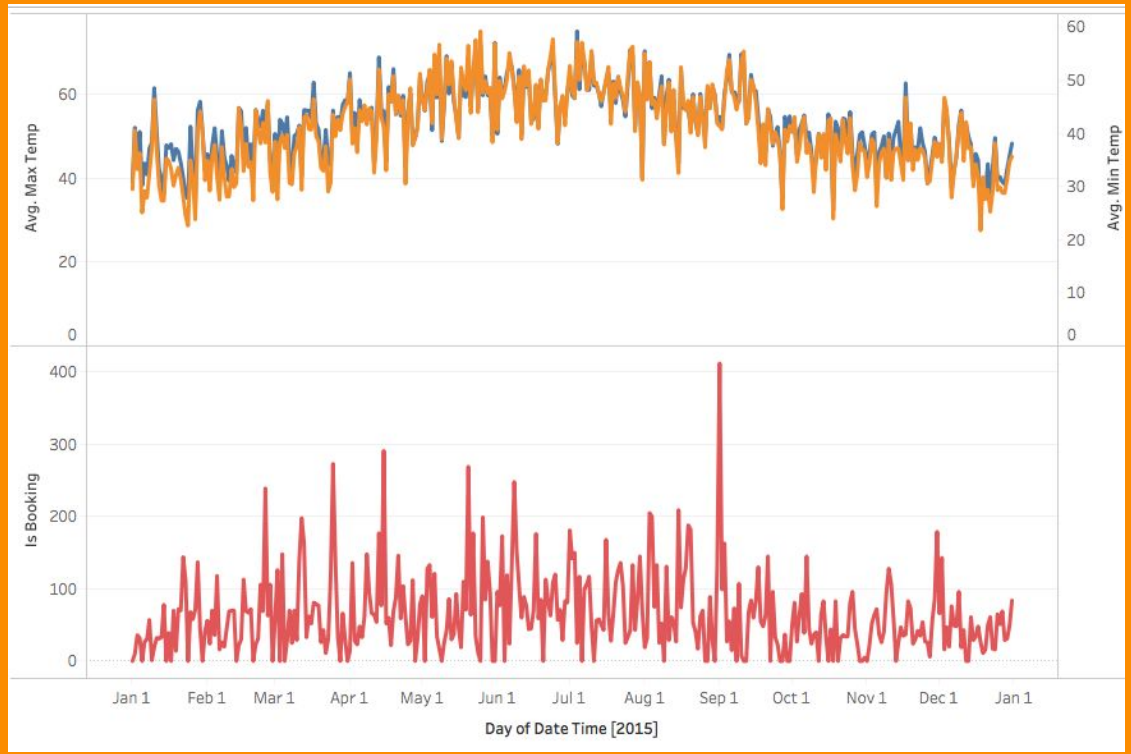
Team Offset

Aarsh Patel
Antonio Lopes Borges
Chenhao Huang
Lynn Samson
Michael Sadler

Data Analysis with Weather and Booking Rates

- We wanted to see if weather affected booking rates
- Extracted external weather metadata using the **checkin_date** and **lat/lng coordinates** for the destination
- Also tried using Machine Learning to figure out which features are important in predicting if a user will book or not:
  - channel
  - cnt



Checkin Dates & Checkout Dates

```
('Checkin Date: ', datetime.datetime(2015, 12, 22, 0, 0), 'Checkout Date: ', datetime.datetime(2015, 12, 16, 0, 0))
('Checkin Date: ', datetime.datetime(2015, 11, 16, 0, 0),'Checkout Date: ', datetime.datetime(2015, 11, 14, 0, 0))
('Checkin Date: ', datetime.datetime(2015, 11, 9, 0, 0),'Checkout Date:', datetime.datetime(2015, 9, 18, 0, 0))
```

- We found that 1755 of these rows had checkout dates before checkin dates
- We also found that lots of these rows had dates that were invalid or didn’t contain any dates

Hit Rates for Booking

- We analyzed the influence of 3 key predictors(international, prop_is_branded, cnt) to determine their impact on the probability that a user books a hotel room(is_booking).
- We found that travelling domestically always has higher hit rates for booking than travelling internationally.
- Also branded always has a higher hit rates for booking than unbranded, and the branded hit rate is nearly equal to domestic hit rate. (as well as international is roughly equal to unbranded).
- Another trend we found, which is somewhat intuitive, is that multiple clicks lead to bookings much less often than single clicks. (by nearly 15x!!!)