

# Project (REPLACE WITH TITLE)

LAM-duh: Xuliang Deng, Leah Okamura, Megan Richards

Nov 15th, 2021

```
library(tidyverse)

## Warning in system("timedatectl", intern = TRUE): running command 'timedatectl'
## had status 1

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.5      v purrr 0.3.4
## v tibble 3.1.5       v dplyr 1.0.7
## v tidyr 1.1.4        v stringr 1.4.0
## v readr 2.0.2        v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(broom)
library(knitr)

hotels <- read_csv("data/hotels_reduced.csv")

## Rows: 2097 Columns: 32

## -- Column specification -----
## Delimiter: ","
## chr (14): hotel, arrival_date_month, meal, country, market_segment, distribu...
## dbl (18): is_canceled, lead_time, arrival_date_year, arrival_date_week_numbe...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

## Introduction and Data

INSTRUCTIONS: This section includes an introduction to the project motivation, data, and research question. Describe the data and definitions of key variables. It should also include some exploratory data analysis. All of the EDA won't fit in the paper, so focus on the EDA for the response variable and a few other interesting variables and relationships.

### Introduction

With the lifting of travel restrictions into the U.S. (<https://www.nytimes.com/2021/09/22/travel/us-international-travel-vaccine.html>) through the implementation of new travel guidelines, we believe that the booking of hotels may start to increase. Therefore, with the slower return to travel and society pre covid, we are interested in studying the characteristics of hotel room reservations in the United States. Specifically, we are interested in what relationship these characteristics have the cost of a hotel. Our general research

question is; How do the characteristics of a hotel booking affect the daily cost of a hotel stay in the United States? We believe there will be several significant points of relevance for understanding these relationships: understanding predictors of room cost could be used to help identify where new hotels could be successfully created, allow travelers to plan financially for future travel.

Generally, we are looking to use linear models to understand the contributing factors to hotel room price, as well as identify the strongest predictors. We hypothesize that a model with predictors of hotel type, reserved room type, assigned room type, company, meal, number of adults/children/babies, the average daily rate or daily cost, and the reservation status, will be statistically significant predictor of hotel room price, and that the predictors will be significant except for company and number of adults/children/babies.