README File
Food Inspection: Modeling Risk Assignment Using Ordinal Logistic Regression
MSCA 31010 Linear and NonLinear Models, Winter 2021
Authors: Christopher Hein, Yuming Liao, Andrew McCurdy, Nina Randorf, Melanie Tran

Description
This README file describes the project background, deliverables, data source files, data preparatory scripts, exploratory and modeling scripts, and the final presentation. This project uses data pulled from the Chicago Data Portal. The City of Chicago voluntarily provides data to its Chicago Data Portal and requires users of the data to include the following disclaimer:

> "This site provides applications using data that has been modified for use from its original source, www.cityofchicago.org, the official website of the City of Chicago. The City of Chicago makes no claims as to the content, accuracy, timeliness, or completeness of any of the data provided at this site. The data provided at this site is subject to change at any time. It is understood that the data provided at this site is being used at one's own risk."

Project Description
This project investigates what factors influence risk assignment for establishments which pass their food inspections using Ordinal Logistic Regression. The data for this project is sourced from *The City of Chicago Data Portal* which hosts information on food inspections from July 2018 to present. The risk assignment is the ordinal response variable (low, med, high) while the season of inspection and health checklist violations are the categorical explanatory variables.

Applications:
The following applications are required to execute and replicate this project.
- Jupyter Notebook, vers 6.0 or later
- RStudio, vers 4.0 or later

Data Source Files:
Delivery Folder: Data
Instructions: Pull the CSV file directly form the Data folder or download as CSV from the linked website. If pulling from the data portal website, subset the data to only include July 2018 to February 19,2021 inspections to replicate results.

File Name	Description
Food_Inspections_-_7_1_2018_-_Present.csv	Raw CSV file pulled from The City of Chicago Data Portal containing all food inspection data from July 2018 to February 19, 2021 (date of pull). Input file for the "Cleaned_Inspections_2018_Forward.ipynb" file. https://data.cityofchicago.org/Health-Human-Services/Food-Inspections-7-1-2018-Present/qizy-d2wf/data
cleanedInspection.csv	Output of the "Cleaned_Inspections_2018_Forward.ipynb" file. Input file for the "OLS Final Project.Rmd" script.

Data Preparatory Scripts:

Delivery Folder: Scripts

Instructions: Open the .ipynb file in Jupyter Notebook or select the .html file to view as read-only. If executing the .ipynb file, the "Food_ Inspections_-_7_1_2018_-_Present.csv" file must be saved in the same working directory as the python script.

File Name	Description
Cleaned_Inspections_2018_ Forward.ipynb	Executable Python script that conducts initial exploratory analysis and cleans a raw CSV file. Takes as input the "Food_ Inspections_-_7_1_2018_-_Present.csv" file (must be saved in the same working directory as the python script). Outputs a cleaned CSV File called "cleanedInspection.csv"
Cleaned_Inspections_2018_ Forward.html	Read only version of the Python script.

Data Modeling Scripts:

Delivery Folder: Scripts

Instructions: Open the .Rmd file in RStudio or .html version to view as read-only.

File Name	Description
OLS Final Project.Rmd	Executable R script to create Ordinal Logistic Regression model. Takes as input the "cleanedInspection.csv" file.
OLS Final Project.html	Read only version of the R script.

Data Dictionary Files:

Delivery Folder: Docs

Instructions: Open the files to view.

File Name	Description
HealthCodeRequirements.pdf	A PDF file outlining the background and health checklist requirements for food inspections conducted and documented by the City of Chicago. https://www.chicago.gov/city/en/depts/cdph/provdrs/healthy_restaurants/svcs/understand_healthcoderequirementsforfoodestablishments.html
Food_Inspection_Data_Description.pdf	A PDF file describing data elements included in the "Food_ Inspections_-_7_1_2018_-_Present.csv" file. The PDF is provided by the City of Chicago Data Portal and is also available to view at the link below. Of note, the data description of the 'violations' is out-of-date; description fo health code checklist 'violations' should be referenced in the "HealthCodeRequirements.pdf" file. https://data.cityofchicago.org/api/assets/BAD5301B-681A-4202-9D25-51B2CAE672FF

Final Project Presentation File:

Delivery Folder: Docs

Instructions: Open the .pdf file to view.

File Name	Description
Ordinal Logistic Regression Final Project.pdf	PDF file of final project presentation.