|  |
| --- |
| **INTRODUCTION** |
| Throughout the fiscal year 2019-2020, the missed appointments or “no-show” data has been recorded at Harlem Hospital Center. Analysis of the data and other variables recorded could potentially give us an insight as to how this information may impact financial sustainability and quality of care given to patients.  Understanding baseline information and trend analysis is a vital component when undergoing quality improvement projects that may impact the overall vision, mission, and strategic goals of the institution and the health system as a whole.  This data set and analysis may affect the NYC Health and Hospital’s strategic pillars (SP) which aligns with Harlem Hospital Centers long term goals for the growth and development of the institution. The specific strategic pillars involved are:   * Improve quality of patient care and services (SP: Quality, Safety) * Increase revenue streams (SP: Financial Sustainability) |
| **DATA SOURCES** |
| **Data Acquisition**  Raw data was acquired from EPIC Reporting Work Bench. Additional data was acquired from multiple relevant websites which includes holidays and weather information.  **Initial Data Set:**   * Data Variables: ‘*mrn', 'dob', 'age', 'sex', 'phone', 'encounter\_provider', 'visit\_date', ‘'appt\_time', 'month', 'day\_of\_the\_week', 'holiday', 'weather\_type', 'max\_temp', 'mini\_temp', 'avg\_temp', 'departure', 'hdd', 'cdd', 'precipitation', 'new\_snow', 'snow\_depth', 'status', 'encounter\_closed', 'dept', 'dept\_id', 'Avg\_payment', 'visit\_type', 'visit\_type\_id', 'copay\_paid', 'chkin\_time', 'check\_in-check out', 'encounter\_diagnoses', 'next\_appt', 'appt\_type', 'patient\_employer', 'employer\_state', 'patient\_state', 'time\_with\_provider', 'wait\_time\_for\_provider', 'zip\_code', 'pcp', 'reason\_for\_visit', 'roomed\_time', 'employer\_zip', 'attending\_provider', 'service', 'language', 'religion', 'hosp\_acct', 'special\_needs'* * Number of rows: 21226 * Number of columns: 50   **Final Data Set:**   * Data Variables: *'mrn', 'dob', 'age', 'sex', 'sex\_id', 'phone', 'encounter\_provider', 'visit\_date', 'appt\_time', 'month', 'year', 'day\_of\_the\_week', 'holiday', 'weather\_type', 'avg\_temp', 'precipitation', 'new\_snow', 'dept', 'dept\_id', 'Avg\_payment', 'visit\_type', 'visit\_type\_id', 'encounter\_diagnoses', 'appt\_type', 'patient\_employer', 'employer\_state', 'patient\_state', 'zip\_code', 'pcp', 'reason\_for\_visit', 'employer\_zip', 'attending\_provider', 'service', 'language', 'religion', 'special\_needs', 'value\_count'* * Number of rows: 21225 * Number of columns: 37 |
| **METHODOLOGIES** |
| * All sensitive data was changed our deleted from the file * Initial data set was in .xlxs file which was converted to .csv file * CSV file was uploaded in Github for easy reference for Python analysis * Google Colab was used for python analysis * Python libraries used: *pandas, numpys, matplotlib.pyplot, matplotlib inline, plotly.express, seaborn, sodapy* * Data cleaning and wrangling was done using Excel and Python   + Dropping unnecessary columns & empty rows   + Adding columns for quantitative analysis   + Analyzing average payments for insurances * Python visualization used: *bar plots, scatterplots, histograms* * Tableau was used to create dashboard and graphs for easy data visualization and presentation * PowerPoint presentation was used for presentation purposes |
| **FINDINGS AND CONCLUSION** |
| * U.S Healthcare system loses over $150 billion dollars in revenue each year due to patients not showing up for their appointments. * No Show also impacts a hospital cost allocation in terms of scheduling staff. * No Show also impacts patient care. Missed appointment can lead to higher acuteness of illness and medical costs overall as patients with treatable illness postpone care until their condition becomes urgent (and ultimately costs more). * Harlem hospital alone lost an estimated $2.3 million dollars for the fiscal year 2020. * For the recently concluded FY, over 21k patients did not show up for their appointment. * Our analysis broke down the FY20 No show data set and gave us some very factual insight. Some of the key findings are:   1. Female patients no show rate is higher than their male counterpart.   2. Higher Temperature and rainy weather are linked to higher no-show rate.   3. Patients between the age of 50-60 have the highest no-show rate.   4. Appointment time between 9am and 10am has the highest no-show rate.   5. Looking at weekdays, Tuesday tend to have the highest no-show rate through the year. |
| **REFERENCES** |
| * https://w2.weather.gov/climate/xmacis.php?wfo=okx * https://en.wikipedia.org/wiki/Public\_holidays\_in\_the\_United\_States * https://www1.ncdc.noaa.gov/pub/data/cdo/documentation/LCD\_documentation.pdf * https://www.ncdc.noaa.gov/cdo-web/datasets#LCD * https://www.timeanddate.com/holidays/us/2019 * https://www.timeanddate.com/holidays/us/ * https://www.ncdc.noaa.gov/cdo-web/datasets#LCD |