Xin Tang

September/2024

Data Presentation and Visualization

DSC640-T301

**Week 3-4 assignment**

**Write Up**

**Story background:** In United States, White house is a popular tourist destination. At the same time, the White House is a place requires highest security all the time so visiting is under monitor of a security force all the time.

**Customer/Audience**: White House security team.

**Purpose**: Do an analysis of visitor logs to determine if they need to increase team size facing increasing visitor volumes.

**Method**: To simplify the data, Excel is used to extract and pre-process the data from 45 individual monthly files and combine them into 3 new excel files, which then be loaded into python. Python data analysis tool is used to generate visual charts. Last, power point is used to generate summary report.

**Medium**: Since the audience is not programmers but familiar with office software. The result will be presented in the form of **powerpoint**.

**Design**: The audience, white house security team, is not a commercial company, but they are data driven decision makers. So basic charts will be the primary method to deliver information.

**Visual**:

* Line chart with time
* Stacked column chart with time
* Pie chart
* Column chart with time
* Step chart with time
* Box chart (distribution chart)

**Ethical Consideration**:

* Data is being reviewed to understand the meaning of each variable. Irrelevant variable is removed, and NA field be removed if it will distort the fact
* Data is open to public, but the analysis is not. So, it needs to conform to all laws and White house security team rules.
* Since the analysis will bring impact to national security decision making, there are risks that need to be known to the security team.
* I made some assumptions when handling NA. For example: any reservation record without actual entry time means no actual visit.
* Since this is white house released data, I assume data is valid and been verified.
* To reduce security risk. The analysis will not be loaded into public domain (like Githhub).