

MIND: Minería de Datos Assignment 4

Description

- El desarrollo de la asignación requiere dos archivos en PDF:
 - Paper: “Data Visualization and Statistical Graphics in Big Data Analysis”
 - Book chapter: “2. Visual Aids for EDA”
- El artículo presenta diversos casos de la vida real donde han utilizado de alguna forma técnicas para Análisis Exploratorio de Datos, además un marco histórico de diversos gráficos, además una revisión literaria actualizada sobre la visualización en procesos de análisis de datos.
- El capítulo hace parte del libro “Hands-On Exploratory Data Analysis with Python” el cual proporciona amplia teoría y práctica relacionada con el análisis de datos. El capítulo compartido presenta diferentes formas de visualizar los datos en la fase de análisis exploratorio, además código fuente para una posible implementación para cada visualización.

Problem

- Las partes que componen cada documento han sido distribuidas entre los estudiantes del curso para desarrollar una corta presentación de la sección/material asignado. No es suficiente con leer la parte asignada, debe leer por completo el documento para ver el contexto y ofrecer una presentación más coherente sobre el tema.
- Además de leer cada documento por completo, usted debe profundizar y presentar a todo el curso de forma clara y con calidad la sección asignada.
- El tiempo máximo de la presentación es de **6 minutos**, durante ese tiempo se espera tener un desarrollo del tema con calidad, donde explique de forma clara la sección asignada, el problema a resolver, la forma de hacerlo, posiblemente hacer uso de scripts en Jupyter, presentar contexto histórico y otros aspectos que brinden profundidad la sección.
- La presentación será el día de clase, viernes 24 de septiembre, comenzamos 1:10pm en punto.
- El objetivo de la presentación NO es copiar, traducir y leer lo plasmado por los autores en los documentos, eso es un ejemplo de mala calidad en una presentación.
- La siguiente tabla detalla las partes que componen en capítulo compartido:

| No. | Section |
|-----|----------------------------|
| 1. | Line and Bar chart |
| 2. | Scatter plot |
| 3. | Area plot and stacked plot |
| 4. | Pie and Table chart |
| 5. | Polar chart and Histogram |
| 6. | Lollipop chart |
| 7. | Choosing the best chart |
| 8. | Other libraries |

- La siguiente tabla detalla cada sección del paper:

| Section | Title |
|---------|--|
| 2.1. | How to Win a Data Mining Challenge |
| 2.2.1. | Organization for Economic... |
| 2.2.2. | Elections |
| 2.2.3. | Airline traffic |
| 2.2.4. | Wikipedia |
| 2.2.5. | The San Francisco housing crisis |
| 2.2.6. | Credit card purchases |
| 3.1.1. | Scatter-plots |
| 3.1.2. | Table plots |
| 3.1.3. | Parallel coordinates |
| 3.1.4. | Rearranging, summarizing, |
| 3.1.5. | Interactive graphics |
| 3.2. | Statistics, Information Visualization... |

- La siguiente tabla relaciona el las iniciales del nombre de cada estudiante y la sección asignada:

| Student | Section | Document |
|-------------|---------|----------|
| ABAQUERO | 2.1. | Paper |
| ACOLORADO | 2.2.1. | Paper |
| CSORZA | 2.2.2. | Paper |
| CTORRES | 2.2.3. | Paper |
| CVIVAS | 2.2.4. | Paper |
| EHENAO | 2.2.5. | Paper |
| GCASALLAS | 2.2.6. | Paper |
| JGALINDO | 3.1.1. | Paper |
| JRODRIGUEZ | 3.1.2. | Paper |
| JVILLARRAGA | 3.1.3. | Paper |
| JFERNANDEZ | 3.1.4. | Paper |
| JPULIDO | 3.1.5. | Paper |
| JFLOREZ | 2.2.2. | Paper |
| JARBELAEZ | 2.2.4. | Paper |
| JGOMEZ | 3.2. | Paper |

| Student | Section | Document |
|------------|----------------------------|--------------|
| LMANUEL | Line and Bar chart | Book chapter |
| LROZO | Scatter plot | Book chapter |
| MMUÑOZ | Area plot and stacked plot | Book chapter |
| NMENDOZA | Pie and Table chart | Book chapter |
| PLOPEZ | Polar chart and Histogram | Book chapter |
| SDIAZ | Lollipop chart | Book chapter |
| SRODRIGUEZ | Choosing the best chart | Book chapter |
| SHERNANDEZ | Choosing the best chart | Book chapter |
| WDELGADO | Other libraries | Book chapter |