

Minería de Datos

Lab. #2

Ivan Saavedra, Ph.D.

saavedrai@uninorte.edu.co

Universidad del Norte
División de Ingenierías
Dpto. Ingeniería de Sistemas



202110

1. Stock Market Analysis

Usted ha sido asignado para desarrollar un Stock Market Analysis usando los datos históricos (últimos 2 años) correspondientes a las acciones bajo estudio.

- Stocks: Apple, Walmart, Royal Caribbean, y Delta. Seleccione una stock de una empresa Colombiana (Ecopetrol, Avianca, Tecnoglass, Bancolombia, etc.)
- Analice las acciones de manera general utilizando los valores de cierre de los dos últimos años. Utilice medidas de tendencia central y de dispersión para su análisis.
- Calcule la media móvil (Moving Average) utilizando los últimos 50 y 100 días. Analice ambos resultados de las media móviles usando las medidas de tendencia central y de dispersión para su análisis.
- Compare los datos históricos vs las medias móviles. Utilice gráficos para esta comparación. Concluya cual es la importancia de las medias móviles en el contexto bajo estudio.
- Cual es la stock con mejor rendimiento? Porque?
- Cual es la stock con menor rendimiento? Porque?
- Encontró algún impacto de la pandemia en las stocks analizadas ? Explique.

Metodología

- **Búsqueda de datos históricos:** Usted necesitara identificar fuentes confiables que suministren la información necesaria para desarrollar este análisis.
- **Organización de los datos:** Usted necesitara organizar los datos de manera apropiada para el Desarrollo del análisis.
- **Análisis de los datos:** Usted necesitara desarrollar un análisis exploratorio para cumplir con el objetivo que ha sido asignado.
- **Visualización de los datos:** Usted necesitara crear múltiples visualizaciones para dar soporte a su análisis exploratorio.
- **Conclusiones:** Usted necesitara resaltar las lecciones aprendidas y conclusiones al final de sus análisis.

2. Vehicle Accident Analysis

Metodología

- **Búsqueda de datos históricos:** Datos históricos de la accidentalidad en la ciudad de Barranquilla han sido descargados de la siguiente fuente :<https://www.datos.gov.co/en/Transporte/Accidentalidad-en-Barranquilla/yb9r-2dsi>. Los datos son suministrados en este ejercicio para su respectivo análisis.
- **Organización de los datos:** De ser necesario, usted deberá organizar los datos de manera apropiada para el desarrollo del análisis.
 - Realice la organización de los datos dentro de su notebook.
- **Análisis de los datos:** Usted necesitara desarrollar un análisis de probabilidad para responder a las preguntas objetivo.
 - Identifique que distribución de probabilidad se ajusta a la variable Cantidad de Accidentes Diarios?
 - Cual es la probabilidad de que se den menos de 10 Accidentes?
 - Cual es la probabilidad de que se presenten mas 15 accidentes?
- **Visualización de los datos:** Usted necesitara crear las visualizaciones para dar soporte a su análisis.
 - Valor esperado vs valor observado
 - Diagrama de cajas
 - Otro?
- **Conclusiones:** Usted necesitara resaltar las conclusiones al final de sus análisis.

Lab-2

Entregable:

- Los archivos con los datos usados para su análisis
- Un archivo de jupyter notebook con el desarrollo del análisis
- Las conclusiones y respuestas al objetivo del análisis deben ser contestadas en el mismo notebook.
- Se sugiere que comente las secciones de manera adecuada para una mejor interpretación de su análisis.
- La fecha de entrega es el Lunes 01 de Marzo del 2021 vía catalogo web enlace de laboratorios.