

# Elementos-Pry3-IISem2022

October 28, 2022

## 0.1 Elementos de Computación

## 0.2 Proyecto No. 3

Construya una notebook Jupyter que sirva como programa/documento explicativo de la realidad actual del proceso de llenado del álbum del mundial Qatar 22. Considere para el documento **celdas de texto y gráficos que en conjunto respondan** como mínimo a las siguientes preguntas: El album tiene la siguiente estructura El álbum tiene la siguiente estructura: (ver <https://www.cardboardconnection.com/2022-panini-world-cup-stickers-qatar-cards>). Para los efectos de este proyecto se omitirá la sección FWC y la postal 00.

1. ¿Cuántas postales han sido obtenidas en las últimas **n** semanas?,Cuál es la distribución porcentual por semana.?

**Nota:** Inicialmente se supone se dispone de 1 semana de datos pero para el tiempo de entrega por lo menos, se espera esten disponibles datos de 2 semanas más Para las siguientes preguntas utilice los datos más recientes disponibles:

2. ¿Cuáles son las 5 secciones que presentan mayor cantidad de postales?, por grupo?.
3. ¿Cuáles son las 5 secciones que presentan menor cantidad de postales?, por grupo?.
4. ¿Cuáles son las 5 secciones que presentan mayor cantidad total de postales repetidas?, por grupo?.
5. ¿Cuáles son las 5 secciones que presentan menor cantidad total de postales repetidas?, por grupo?.
6. ¿Cual es la distribución porcentual de postales (la tengo, no la tengo, la tengo repetida) para un grupo dado?. Se debe leer el grupo.
7. ¿Cuál es la distribución porcentual de las postales (la tengo, no la tengo, la tengo repetida) por sección?. Se debe leer la sección
8. ¿Cuál es la distribución porcentual de las secciones (la tengo, no la tengo, la tengo repetida) con respecto al escudo (postal 1 de cada sección)?
9. ¿Cuál es la distribución porcentual de las secciones (la tengo, no la tengo, las tengo repetidas) con respecto a sus porteros (postales 2 y 3 de cada sección)?
10. ¿Se cuenta con las postales de Lionel Messi, Cristiano Ronaldo, Mbappe, Karim Benzema, Neymar?.

## 1 Aspectos administrativos

1. La entrega es electrónica a más tardar el viernes 11 de Noviembre antes de la medianoche. La entrega se hará en Teams en el canal de Evaluaciones, opción Files, carpeta Proyecto3.
2. Se puede confeccionar y entregar en forma individual o en parejas.

3. El plagio o trabajo ‘colaborativo’ no es permitido y será fuertemente sancionado fuera de la tarea.
4. Dentro de la calificación se tomará en cuenta:
  - Funcionamiento del programa 75%
  - Documentación interna 10%
  - Estética y claridad 15%
5. La respuesta a cada pregunta debe estar claramente redactada y debe explicar el o los gráficos de apoyo a esa respuesta. Es decir, no se considera válida una respuesta que solo incluya o solo texto o solo un gráfico.
6. Pueden agregarse preguntas y respuestas adicionales a las básicas por puntos adicionales. Sin embargo, deben estar apoyadas en datos “reales”.

## 2 Aspectos técnicos

0. Construya archivos .csv sea a partir de los proyectos anteriores o creadas manualmente pero simulando la salida de otro programa.
1. El Jupyter debe estar estructurado como un documento formal, con su introducción y conclusiones, y no limitarse únicamente a presentar respuestas a las preguntas de arriba. Es decir, el documento debe tener un hilo conductor.
2. Se debe incluir abundante documentación interna (comentarios) independientemente de las celdas de texto en el notebook, de acuerdo a la guía suministrada.
3. Debe utilizar los paquetes Pandas, Matplotlib y Seaborn, aunque opcionalmente puede también utilizar otros paquetes de visualización.
4. Escoja el mejor tipo de gráfico en cada uno de los casos dependiendo del tipo de información a publicar.
5. Todos los gráficos deben incluir su título y la referencia a la fecha de los datos mostrados.
6. En varios casos, para obtener las respuestas requieren manipulación y preprocesamiento de datos, todo lo cual debe de hacerse dentro del programa y no modificando los archivos .csv
7. En el caso de la información no disponible en .csv, se deberán crear dataframes a partir de archivos creados manualmente o por otro medio de creación de dataframes. Sin embargo, se sugiere evitar datos “hardwired”.