



UNIVERSIDAD POLITÉCNICA DE CHIAPAS

Propuesta de proyecto:

PERSONAS QUE HAYAN VISTO LA SAGA DE **STAR WARS**

Probabilidad y Estadística

Alumno: Ada Isabel Gamboa Chanona

Docente: Doc. Carlos Alberto Díaz Hernández

CUADRO DE IDENTIFICACIÓN

CUATRIMESTRE	Mayo-Agosto
CUATRIMESTRE Y GRUPO	6to "A"
ASIGNATURA	Probabilidad y Estadística
ACTIVIDAD	PE.C1.A1 Propuesta de proyecto
FECHA DE ASIGNACIÓN	Jueves, Mayo/13/2021
FECHA DE ENTREGA	Jueves, Mayo/13/2021
MATRÍCULA	193223
NOMBRE	Ada Isabel Gamboa Chanona

CONTENIDO

CUADRO DE IDENTIFICACIÓN.....	2
LISTA DE FIGURAS.....	3
1. INTRODUCCIÓN.....	4
1.1 Descripción del problema.....	4
2. MÉTODOS	5
2.2 Selección de variables	5
2.2.1 Variable Cualitativa.....	5
2.2.2 Variable Cuantitativa Discreta	5
2.2.3 Variable cuantitativa Continua	5
2.3 Obtención de los datos	5
2.4 Análisis de datos.....	5
2.5 Presentación de datos	6
3. IMPLEMENTACIÓN	7
4. BIBLIOGRAFÍA.....	8

LISTA DE FIGURAS

Figura 1. Ejemplo de Polígono de Frecuencia	6
Figura 2. Ejemplo de Gráfica de Pastel	6
Figura 3. Ejemplo de Diagrama de Barras	6
Figura 4. Python	7
Figura 5. Numpy	7
Figura 6. Pandas	7
Figura 7. Matplotlib	7

1. INTRODUCCIÓN

La estadística es una ciencia y una rama de las matemáticas a través de la cual se recolecta, analiza, describe y estudia una serie de datos a fin de establecer comparaciones o variabilidades que permitan comprender un fenómeno en particular.

Hay varios tipos de estadística, pero en este proyecto utilizaremos la estadística descriptiva, que nos permite presentar de manera resumida y organizada los datos numéricos obtenidos tras un estudio o análisis en particular. Su objetivo, por lo tanto, es describir las características principales de los datos reunidos y evitar generalizaciones.

En el proceso que seguiremos para llevar a cabo este tipo de estadística, será como primeros pasos la selección del tema, la obtención de los datos y el análisis de los mismo, presentándolos en diagramas prácticos como tablas y gráficas

El tema a tratar es “Las personas que han visto la saga de **STAR WARS**” donde averiguaremos datos como la película favorita de la persona, la cantidad de películas que ha visto y el tiempo que ha tomado viéndose las películas o la saga completa. Con ello podremos evaluar, por ejemplo, la película más popular, si las personas en general les gusta esta saga y el fanatismo que presentan conforme el tiempo que invierten viendo la franquicia.

1.1 Descripción del problema

La saga de películas de STAR WARS es una franquicia que se presentó en pantalla desde 1977, siendo su primera película ‘STAR WARS: Una nueva esperanza’ después del éxito de la película empezó a salir una saga completa y en la actualidad la franquicia tiene 9 películas principales sin contar las series y demás películas que cultivan la historia, debido a la cantidad de películas, resulta algo tedioso y muy tardado ver toda la saga si en realidad no es de tu gusto, del mismo modo, a las personas pueden resultarles ligeras y entretenidas si son verdaderos fans. Por ello, esta investigación busca obtener a las personas que han visto la saga y determinar el gusto por la franquicia o una de las películas, el tiempo que invierten en verla y la cantidad de películas que han visto, utilizando preguntas que al final serán presentadas en gráficas para analizar los resultados.

2. MÉTODOS

2.2 Selección de variables

Para seleccionar las variables que se evaluarán se han establecido criterios sobre los requerimientos del proyecto es por ello, que se establecieron tres variables principales para abarcar el tema

2.2.1 Variable Cualitativa

La variable cualitativa que se establecerá será '*¿Cuál es tu película favorita de la saga?*' tomando como respuesta cada una de las películas principales de la franquicia, la variable se identifica como cualitativa porque no representan a un número o valor contable.

2.2.2 Variable Cuantitativa Discreta

La variable cuantitativa discreta se representará con la cantidad de películas que ha visto el encuestado, es discreta debido a que la cantidad de películas se reduce a nueve, por lo tanto, es un valor entero y definido.

2.2.3 Variable cuantitativa Continua

Para la variable cuantitativa continua se preguntará el tiempo que ha invertido viendo las películas, siendo continua debido a que el individuo puede haber invertido 0 horas (en el caso de que no haya visto las películas o 1 hora, 1 hora y media, etc. En caso de que haya visto la mitad de una película, una película completa o haya hecho un maratón de la saga.

2.3 Obtención de los datos

Para obtener los datos necesarios se usará la herramienta Google Forms, en donde se realizará la encuesta para posteriormente compartirla a 200 personas en redes sociales (ya sea WhatsApp o Facebook) para que sea contestada y así poder analizar los datos obtenidos, adquiriendo así un tipo de muestra no probabilística, que, en este caso, es la manera más rápida y efectiva para recabar la información.

2.4 Análisis de datos

Para el análisis de datos se recogerán los obtenidos en la encuesta y se analizarán cada uno de ellos utilizando las medidas de tendencia central. En los datos cuantitativos, se sacará la media aritmética para obtener el promedio de los datos, la media geométrica y la media

truncada con un porcentaje de 10% con la finalidad de analizar cada uno de los resultados para cálculos posteriores.

Del mismo modo se calculará la mediana de los datos cuantitativos, con la finalidad de obtener el dato central de la serie de datos, y se realizará el cálculo de la moda, para corroborar la frecuencia más alta extraída.

También recabaremos otras características con el análisis de los datos, tales como el sesgo de la gráfica que se presentará, la varianza muestral, la desviación estándar y del mismo modo la distribución de frecuencia de los datos cuantitativos continuos, tomando en cuenta conceptos como el rango, el número de clases, ancho de la clase, limites, marca de clase, y la distribución de frecuencia absoluta y relativa.

2.5 Presentación de datos

Para presentar los datos del análisis se usará la herramienta Excel donde se almacenarán los datos obtenidos en diferentes tablas.

Para las gráficas se utilizarán diferente presentación para cada variable, se dispondrá de una *gráfica de polígono de frecuencia* para los datos continuos como la que aparece en la Figura 1. Para los datos discretos, se especificarán en una *gráfica de pastel* (Figura 2) en donde se sacará la frecuencia relativa de los datos para poder realizarla, y para la variable cualitativa se representarán los datos en un *diagrama de barras* (Figura 3)

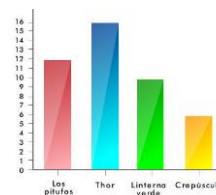
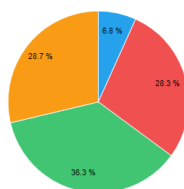


Figura 1. Ejemplo de Polígono de Frecuencia Figura 2. Ejemplo de Gráfica de Pastel Figura 3. Ejemplo de Diagrama de Barras

3. IMPLEMENTACIÓN

Para la implementación del proyecto y enfocarlo en el desarrollo de software, se realizará un programa en Python utilizando Ciencia de Datos para llevar a cabo todas las operaciones estadísticas del proyecto y del mismo modo que sea capaz de realizar las gráficas con solo meter un archivo .csv. El programa tendrá como característica que dispondrá de una interfaz visual para facilitar el uso de la aplicación.

Para realizarlo, se usarán diferentes métodos y librerías de Python que ayudarán a automatizar el proceso de programación, así como también nos brindarán todos los recursos necesarios, ya que estas librerías contienen características específicas que se pueden aplicar a los cálculos estadísticos, por ejemplo:

- Numpy: Es una biblioteca para el lenguaje de programación Python que da soporte para crear vectores y matrices grandes multidimensionales, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas.
- Pandas: Es una biblioteca de Python que sirve para tener un objeto DataFrame rápido y eficiente para la manipulación de datos con indexación integrada. Del mismo modo es una excelente herramienta para leer y escribir datos entre estructuras de datos en memoria y diferentes formatos: archivos de texto y CSV, Microsoft Excel, bases de datos SQL y el rápido formato HDF5;
- Matplotlib: Es una biblioteca completa para crear visualizaciones estáticas, animadas e interactivas en Python. Sirve para desarrollar gráficos de calidad y figuras interactivas que puedan hacer zoom, desplazarse, actualizar etc.

En general, estas son las herramientas que se utilizarán para la elaboración del programa y con las cuales harán que funcione correctamente.



Figura 4. Python



Figura 5. Numpy



Figura 6. Pandas

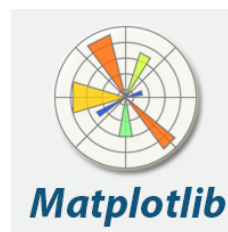


Figura 7. Matplotlib

4. BIBLIOGRAFÍA

Numpy.org. 2021. What is NumPy? — NumPy v1.20 Manual. [online] Available at: <<https://numpy.org/doc/stable/user/whatisnumpy.html>> [Accessed 13 May 2021].

Significados. 2021. Significado de Estadística. [online] Available at: <<https://www.significados.com/estadistica/>> [Accessed 13 May 2021].

Fandom, S., 2021. Star Wars. [online] Star Wars Wiki. Available at: <https://starwars.fandom.com/es/wiki/Star_Wars#:~:text=Logotipo%20de%20Star%20Wars.&text=Star%20Wars%2C%20traducido%20al%20espa%C3%B1ol,ficci%C3%B3n%20creado%20por%20George%20Lucas.> [Accessed 13 May 2021].

Pandas.pydata.org. 2021. pandas - Python Data Analysis Library. [online] Available at: <<https://pandas.pydata.org/about/index.html>> [Accessed 13 May 2021].

Matplotlib.org. 2021. Matplotlib: Python plotting — Matplotlib 3.4.2 documentation. [online] Available at: <<https://matplotlib.org/>> [Accessed 13 May 2021].