



# Bitácora

## Verano de la ciencia 2022.

**Proyecto:** “Aplicación de machine learning para evaluar el impacto de covid-19 en el ámbito educativo.”

Arturo Daniel Jimenez Salazar

Fecha	Actividades	Recursos	Comentarios
23/06/2022	Inauguración del Verano de la ciencia	Instalaciones de la UG	Lamentablemente no pude asistir en la sede del evento en Guanajuato, pero el proyecto tuvo inicio éste día
27/06/2022 a 30/06/2022	<ul style="list-style-type: none"><li>• Investigación de antecedentes y estudios previos al tema de investigación</li><li>• Recopilación de artículos y bases de datos referentes al tema de investigación</li><li>• Preparación de diapositivas y exposición de los resultados, métodos y datos encontrados tras la investigación.</li></ul>	Conexión a Internet, laptop, MS powerpoint.	Se encontraron muchos estudios realizados en cuanto a evaluaciones psicológicas tras la pandemia y el rendimiento educativo, resaltando dos en los cuales basé mi primera investigación de antecedentes, el cual fue realizado en estudiantes de Arabia Saudita, Jordania y Egipto y tenía como peculiaridad que analizaba ciertos factores con el fin de ver que tan relacionados estaban con el rendimiento académico, razón por la que lo elegí sobre otros.
01/07/2022	Presentación de la información encontrada de los artículos recopilados de la investigación de antecedentes	Conexión a Internet, laptop, MS powerpoint, MS Teams.	Se expuso un panorama general de donde estábamos y a donde queríamos llegar, por lo que fue un buen punto de partida para comenzar el estudio.
01/07/2022 a 08/07/2022	Revisión y estudio del libro Python Data Science Handbook (Jake VanderPlas) para contextualizar el uso de las herramientas a utilizar durante el manejo de los datos	Conexión a Internet, laptop.	Se estudió en este periodo el lenguaje de programación Python a profundidad así como diferentes algoritmos de Machine Learning tales como los mencionados en los artículos que se estudiaron.
08/07/2022	Reunión entre los integrantes del verano con el fin de organizar las	Conexión a Internet, laptop, MS Teams.	Como parte de las actividades acordadas nos reunimos para darle seguimiento a la forma en

	actividades para la recaudación de datos		que recabaríamos los datos de la investigación.
08/07/2022	<ul style="list-style-type: none"> <li>• Traducción y adaptación de la encuesta utilizada en uno de los artículos previamente analizados con el fin de tener una base sólida previamente documentada de los parámetros a estudiar</li> <li>• Formulación de preguntas de carácter innovativo sobre los parámetros a estudiar y darle el enfoque buscado según los objetivos de la investigación</li> <li>• Creación de la primera versión del documento de Google forms que contiene las preguntas ya formuladas anteriormente con el fin hacer una encuesta para la comunidad universitaria y así recabar los datos necesarios para la investigación</li> </ul>	Conexión a Internet, laptop, MS Teams, Google Forms.	Después de descartar numerosas encuestas y escalas psicológicas nos decantamos por uno de los artículos que estudié inicialmente puesto que contaban con la encuesta tal como se aplicó (con su traducción al inglés) por lo cual había que modificar de tal forma que nos sirviera para el contexto en el que trabajaríamos, que es el de la Universidad de Guanajuato.
11/07/2022	Revisión de las preguntas para la encuesta a realizar y adecuación de las preguntas	Conexión a Internet, laptop, MS Teams, Google Forms.	Tuvimos la primera revisión de la encuesta realizada con el fin de verificar el enfoque y que tan acertadas eran las preguntas para lo que queríamos estudiar.
12/07/2022	Adición de preguntas relacionadas al uso de lenguaje automático	Conexión a Internet, laptop, MS Teams, Google Forms.	Se tuvo la segunda reunión con el fin de dar ya el enfoque del uso del lenguaje automático al rumbo de la investigación con el fin de dar un carácter innovador al estudio realizado en contraste a lo que ya se había hecho.
15/07/2022	Adecuación final de la encuesta que se enviará a los estudiantes	Conexión a Internet, laptop, MS Teams, Google Forms.	Se hicieron los ajustes finales para la difusión de la encuesta.
16/07/2022	<ul style="list-style-type: none"> <li>• Difusión de la prueba piloto de la encuesta realizada para la detección de errores</li> <li>• Una vez corregidos los errores suscitados se hizo la difusión de la encuesta a</li> </ul>	Conexión a Internet, laptop, MS Teams, Google Forms, whatsapp y facebook.	Se tuvieron que ajustar ciertos detalles técnicos pero éste día se comenzó a mandar la encuesta a los estudiantes de la UG.

	través de redes sociales tales como whatsapp y facebook		
18/07/2022	<p>Actividades de difusión física, tales como carteles o banners además de dar seguimiento a la difusión digital:</p> <ul style="list-style-type: none"> <li>• Colocar carteles con un QR para que fuese sencillo escanear y contestar dentro de la división de ciencias e ingenierías campus León</li> <li>• Visita con los carteles a sede Forum y San Carlos</li> </ul>	Google Forms, whatsapp y Facebook, instalaciones de sede Campestre, Forum y San Carlos	La sede con mejor acceso fue DCI debido a las diferentes actividades que tienen mayor flujo de estudiantes a comparación de San Carlos o Forum
19/07/2022	<ul style="list-style-type: none"> <li>• Capacitación en procesamiento y análisis de textos en ML</li> <li>• Inicio de Análisis de resultados</li> </ul>	Conexión a Internet, laptop, MS Teams, Anaconda Navigator (Python).	Se tuvo una capacitación muy fructífera con el fin de comenzar a preprocesar los datos obtenidos y una vez llegada la fecha de corte de las encuestas se pudiesen analizar directamente.
21/07/2022	Sesión para análisis y revisión de resultados de la encuesta realizada	Conexión a Internet, laptop, MS Teams, Google Forms.	Se comenzó con en análisis e interpretación de los datos, ya cortando la fecha para responder en dicha encuesta
22/07/2022	<ul style="list-style-type: none"> <li>• Análisis y aplicación de técnicas para procesar los datos de la parte de información demográfica</li> <li>• Descripción e interpretación de resultados</li> <li>• Realización de Nube de palabras para la sección.</li> </ul>	Conexión a Internet, laptop, MS Teams, Google Forms.	Se interpretaron los datos con las herramientas aprendidas para la sección mencionada.
23/07/2022	Revisión de reportes y entregables	Conexión a Internet, laptop, MS Teams, Google Forms.	Se termino de reportar el informe y darle toques finales a las actividades realizadas.

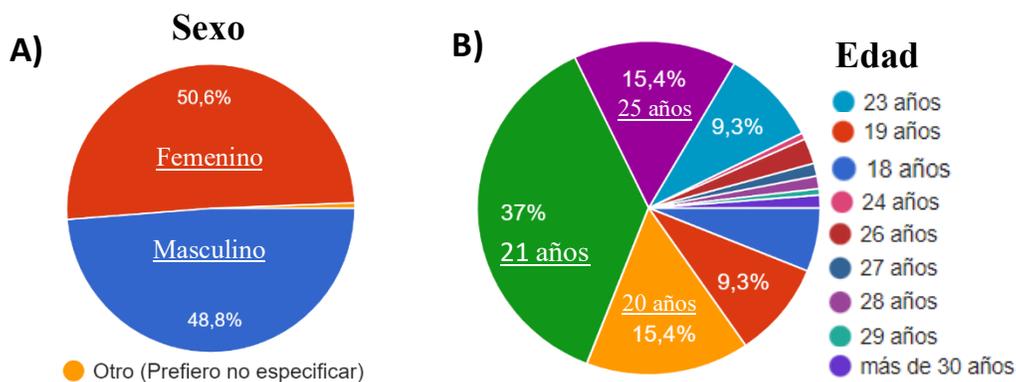
# Reporte de Análisis de datos

## Información demográfica.

En el apartado de información demográfica se preguntó por el sexo del encuestado, edad, licenciatura, semestre que cursa (tomando en cuenta que la encuesta solo se realizó a estudiantes universitarios pertenecientes a la universidad de Guanajuato) y los promedios general y semestral (del último periodo cursado).

Fueron, a la fecha de corte programada para la encuesta, 162 repuestas de las cuales se obtuvo una proporción equitativa entre hombres y mujeres, ligeramente mayor la de mujeres que de hombres, con 82 y 79 respuestas respectivamente, contando solo con una persona que respondió a la opción de no especificar sexo [Fig 1.0].

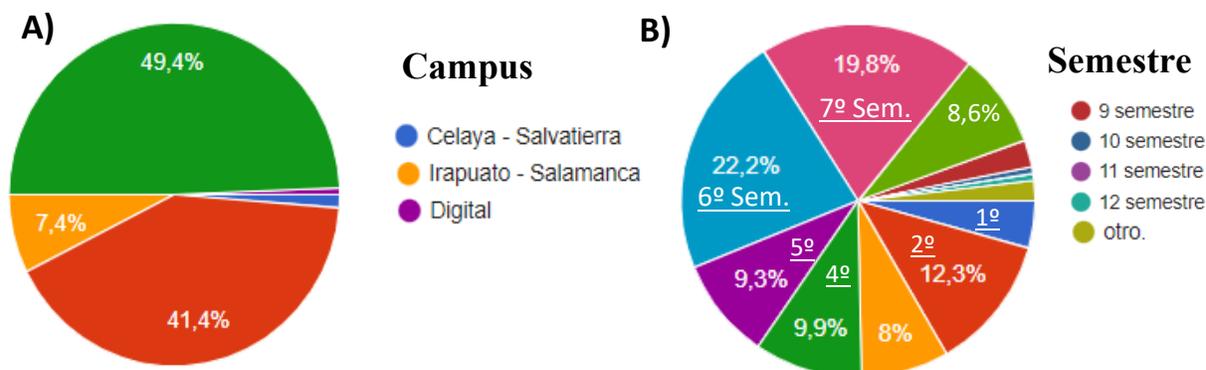
En cuanto a las edades de los encuestados, hubo una **mayoría correspondiente a la edad 21 años con un 37 %** del total (60 repuestas, siguiéndole con un empate entre las personas con **22 y 20 años, con un 15.4%** (25 respuestas) de participación en cada edad, se tuvo después un 9.3% (15 respuestas) tanto en la edad de 23 como en la de 19 años, un 6.8 % para la gente de 18 años, el resto de respuestas (correspondiendo al 7.3 %) oscila entre 1 a 4 respuestas para las edades de 18, 24, 26, 27, 28, 29 y más de 30 [Fig 1.0]. Concentrandose la mayoría de la población entre los 20 y 22 años siendo casi el 70 % de los encuestados.



**Fig. 1.0 A)** Gráfico tipo pastel con porcentajes de la distribución de respuestas de la pregunta referente al sexo del encuestado **B)** Gráfico tipo pastel con porcentajes de la distribución de respuestas de la pregunta referente a la edad del encuestado

De los 50 campus de la universidad de Guanajuato, el campus **con mayor participación fue el campus León** con 80 respuestas de 162 siendo el **49.4%** del total, **el segundo campus con mayor cantidad de respuestas fue el campus Guanajuato** con 67 respuestas, equivalente al **41.4%**, por su parte, 12 personas respondieron del campus **Irapuato – Salamanca (7.4 %)**, 2 personas del campus celaya salvatierra (1.2%) y finalmente solo una persona respondió ser parte del campus digital UG (0.6%) [Fig 1.1]. Podemos ver que si bien,

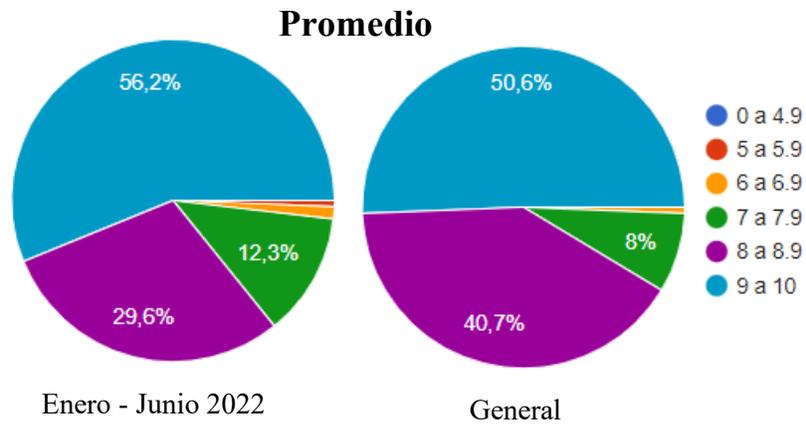
hubo bastante participación de campus guanajuato la mayoría de las respuestas se concentraron en estudiantes del campus león, esto tiene sentido puesto que fue la gente más próxima geográficamente a la que se tuvo acceso hablando de instalaciones y a la que se le pudo hacer mayor difusión, pero que gracias a las redes sociales, pudo extenderse la difusión de la encuesta realizada en todos los campus, en poca u alta proporción, viendo el notable efecto en las respuestas de gente de campus Guanajuato, éstos resultados se verán directamente reflejados a continuación en cuanto a la distribución de los programas educativos reflejados en las respuestas.



**Fig. 1.1 A)** Gráfico tipo pastel con porcentajes de la distribución de respuestas de la pregunta referente al Campus del encuestado **B)** Gráfico tipo pastel con porcentajes de la distribución de respuestas de la pregunta referente al número de semestre que justo finalizó el encuestado (tomando en cuenta el ciclo Enero-Junio 2022)

Respecto a las 72 licenciaturas ofrecidas en los campus anteriormente mencionados se tuvo la participación de 38 programas académicos, donde el **19.8 % de los encuestados** pertenece al programa académico de **Física (Campus León)** con 32 respuestas en la encuesta, **le sigue con un 11.7 %** en mayoría, **el programa académico de Médico cirujano (Campus León)** con 19 respuestas, **en el tercer lugar** de participación se tuvo a la gente del programa académico de **Derecho (Campus Guanajuato) con el 6.2%**, que corresponde a 10 respuestas, dicho esto en cuanto a la mayoría de respuestas, sin embargo hubieron respuestas de estudiantes de Ingeniería física (9 respuestas), Ingeniería Biomédica (8), Biología experimental (7), ingeniería química sustentable (6), Gestión Empresarial (5), Historia (5), Letras españolas (5), Ingeniería Química (5), Contador público (5), Artes Digitales (4), Educación (4), Psicología (4), Matemáticas (3), Química (3), Arquitectura (3), Comercio Internacional (3), Artes visuales (2), Economía (2), Relaciones Industriales (2), Ciencia Política (1), Enseñanza del Español como segunda Lengua (1), Filosofía (1), Sociología(1), Ingeniería en Biotecnología (1), Ingeniería en comunicaciones y electrónica (1), Ingeniería en Minas (1), Ingeniería Mectrónica (1), Diseño de interiores (1), Música (1), Medicina veterinaria y zootecnia (1), Ciencias de la actividad física y salud (1), Nutrición (1), Fisioterapia (1), Administración de la calidad y productividad (1) y Mercadotecnia(1) [Fig 1.2]. Como se mencionó anteriormente, la mayoría de las carreras mencionadas pertenecen a





**Fig. 1.3 )** Comparativo de gráficos tipo pastel con porcentajes de frecuencia con el fin de contrastar los promedios del semestre Enero-Junio 2022 con el promedio general que en principio parte de él corresponde al tiempo en el que se tomaron clases a distancia.

Hablar de rendimiento educativo basándonos únicamente en los promedios de los estudiantes nos daría una visión bastante sesgada del verdadero aprendizaje y rendimiento que tuvieron puesto que hay mucho factores a considerar más allá de alguna calificación o puntuación, razón por la cual se buscó evaluar dicho rendimiento e impacto de la pandemia en el mismo a través de lo obtenido de las respuestas de cada sección de la encuesta.