**ANÁLISIS DE DATOS**

**TDSD**

|  |  |
| --- | --- |
|  | |
| ASIGNATURA: | ANÁLISIS DE DATOS |
| PROFESOR:  FECHA: | Ing. Lorena Chulde / Ing. Juan Pablo Zaldumbide  13 – 08 - 2024 |
| PERÍODO ACADÉMICO: | 2024-A |
|  | |

**PPROYECTO FINAL – BIMESTRE 2**

|  |
| --- |
|  |
| herramientas de análisis de datos y su importancia | UdeCataluña |
|  |
| **Integrantes:**  **Xxxxxxxx**  **xxxxxxxxx** |
|  |

**2024-A**

**Objetivo:**

Aplicar los conocimientos adquiridos sobre análisis de datos, extracción, limpieza, transformación, visualización de datos mediante la aplicación Power BI.

**Proyecto Final Análisis de Datos**

**Indicaciones:**

35% nota final

Fecha de entrega: 13 de agosto, defensa individual.

**Diseñar** una arquitectura en la cual confluyan al menos 10 fuentes de datos de al menos 5 de las siguientes temáticas.

1. Tráfico vehicular en las 5 principales ciudades del Ecuador.
2. Eventos deportivos a nivel mundial.
3. Pulso político en 5 ciudades de Ecuador.
4. Actividades y hobbies.
5. Conciertos y eventos públicos.
6. Tema definido por el grupo.
7. Restaurantes y sitios de esparcimiento.
8. Eventos o noticias mundiales.

**Requisitos:**

* Se debe utilizar al menos dos bases de datos relacionales en su arquitectura (SQL Server, SQLite, MySQL, Oracle, PostgreSQL, etc.)
* Se debe utilizar al menos dos bases de datos NoSQL en su arquitectura (CouchDB, MongoDB, Riak, RavenDB, Redis)
* Se debe evidenciar el paso de datos de una base relacional a NoSQL y viceversa.
* Se debe evidenciar el paso de diferentes tipos de datos entre sí, por ejemplo, json a csv, csv a json, etc.
* El repositorio final debe ser SQL Server

Flecha

Descripción generada automáticamente con confianza baja

**Realizar:**

Crear dashboards explicativos de la información generada. Estos dashboards me permitirán tomar decisiones a problemáticas planteadas por cada uno de los grupos, se deben plantear al menos 15 casos de estudio.

Concluir sobre las visualizaciones y generar casos de estudio.

**Requisitos:**

Sumando los registros de los datasets escogidos, deben tener al menos un millón de registros.

Se debe trabajar con datos locales y en la nube.

**Entregables:**

# Informe (pdf, IEEE doble columna) (25%)

El informe debe contener al menos:

* Definición del caso de estudio.
* Objetivos General y específicos.
* Descripción del equipo de trabajo y actividades realizadas por cada uno.
* Cronograma de actividades incluído el diagrama de gantt (project)
* Recurso y herramientas utilizadas.
* Arquitectura de la solución.
* Extracción de datos.
* Análisis de información.
* Visualización de información.
* Resultados obtenidos.
* Conclusiones y Recomendaciones.
* Desafíos y problemas encontrados.
* Link de github del proyecto.

Se calificará aspectos tales como:

* Calidad del contenido.
* Cantidad de datos recopilados.
* Calidad del análisis.

**2. Videos explicativos (20%)**

* Un video explicativo en YouTube, el cual debe tener al menos 50 vistas, sobre el proceso completo de recopilación de datos. El video debe ser de 4 a minutos.
* Un segundo video explicativo en YouTube, el cual debe tener al menos 50 vistas, sobre las conclusiones de los dashboards. El video debe ser de 4 a minutos.

# Presentación del proyecto (ppt, pdf o prezi) (30%)

* Duración: mínimo de 20 minutos y debe contener los elementos técnicos y los mencionados en el informe.
* Preguntas: Cada asistente debe formular como mínimo una pregunta a un miembro del grupo expositor. Dicha pregunta será calificada tanto al que pregunta como al expositor.

**Nota:** la defensa se calificará de forma individual.

**3. URL del Repositorio de GitHub (25%)**

* Repositorio de GitHub con todos los script utilizados y explicando el proceso completo y las fuentes de datos.

**Nota:** Se tomará en cuenta la innovación en cualquiera de las etapas del proyecto, es un plus para la nota final.