

TRABAJO GRUPAL PRÁCTICO FINAL

Kevin Díaz Cantillano, kevin.diaz@alumnos.uv.cl

Ignacio Salinas, ignacio.salinasa@alumnos.uv.cl

Benjamín Zárate Chacana, benjamin.zarate@alumnos.uv.cl

Sofía Lopez sofia.lopezo@alumnos.uv.cl

Benjamin Morales benjamin.moralesc@alumnos.uv.cl

1. Introducción

Miles de estudiantes eligen su carrera soñada cada año, pero

2. Materiales y Métodos

2.1. Materiales

Los datasets utilizados fueron extraídos desde el sitio oficial del Consejo Nacional de Educación (CNEC), que consiste en [] y pone a disposición diferentes dataset sobre [] de entre los cuales se rescato el dataset [] que contiene información sobre [] entre los años [] y [] este dataset fue elegido para la realización de este trabajo porque [].

¿que dtaset se usaron?, se uso excel, que lenguajes de programacion se usaron, en que consiste, cuales gestores de base de datos se usaron, que herramienta se uso como power bi

2.2. Métodos

Desarrollar un dashboard en Power BI, orientado a un análisis profundo y especializado de cohortes estudiantiles, optimizado para un equipo de Analistas de Datos, que permita analizar el rendimiento y progreso de cohortes en función de año de ingreso, carrera, género y origen. Este dashboard no solo integrará filtros avanzados y gráficos interactivos (como mapas de calor, gráficos de barras y líneas de tiempo) para explorar en profundidad las tendencias de retención y éxito académico, sino que también permitirá desgloses específicos que revelen patrones críticos entre cohortes y segmentos. Al adoptar un enfoque estratégico y detallado, el equipo podrá traducir estos hallazgos en decisiones informadas que fortalezcan las políticas de retención y el éxito académico, aprovechando su experiencia para maximizar el valor de los datos en el diseño de intervenciones y mejoras educativas efectivas.

1. Propuesta de Valor

- Identificación de patrones de retención y éxito académico.
- Visualizaciones interactivas para análisis profundo.

2. Público objetivo, roles de usuario

- Analistas de datos en instituciones educativas.
- Gestores académicos y responsables de retención.
- Analista Senior: Explora datos y patrones avanzados de retención.
- Gestor Académico: Monitorea métricas de éxito y efectúa intervenciones.
- Equipo de Retención: Implementa políticas basadas en insights del dashboard.

3. Fuente de datos

- Base de Datos: BaseDefinitivaINDICES-2005-2024 del CNEC.

Cobertura Temporal: Desde 2010 hasta 2024.

- Variables Incluidas: Matrículas, vacantes, aranceles, puntajes PSU.
- Desglose de Datos: Por institución, región, sede, y género.
- Acceso: Disponible para descarga en <https://cned.leADERA.cl/institucional/bases-de-datos/>.
- Consideraciones: Actualización de datos y consistencia para análisis en Power BI.
- Filtros: Filtros avanzados por año, carrera, género y origen.

4. Preparación y modelado de datos

- **Extracción de Datos:** Descargar y verificar la integridad de la BaseDefinitivaINDICES-200-2024.
- **Limpieza de Datos:** Eliminar duplicados, corregir valores faltantes y errores tipográficos.
- **Transformación:** Estandarizar formatos (fechas, nombres) y normalizar variables clave.
- **Modelado:** Crear relaciones entre tablas (e.g., estudiantes, programas, años).
- **Segmentación:** Definir segmentos por año de ingreso, carrera, género y origen.
- **Carga en Power BI:** Importar los datos transformados en el modelo de Power BI.
- **Optimización:** Configurar columnas calculadas y medidas para análisis eficiente.

5. Almacenamiento del sistema de Backend

- Base de Datos: MySQL para almacenar datos de cohortes, rendimiento y retención.
- Escalabilidad: Estructura flexible para adaptarse al crecimiento de datos.
- Seguridad: Cifrado básico y acceso controlado a los datos.
- Integración con DataMart: Carga directa al DataMart para análisis en Power BI.

6. Modelo de datos Objetivos

- **Cohorte Estudiantil:** Datos de ingreso, carrera, género y origen del estudiante.
- **Desempeño Académico:** Variables de rendimiento como promedio de notas y créditos aprobados.
- **Retención y Abandono:** Indicadores de continuidad o deserción por período académico.
- **Segmentación Temporal:** Año de ingreso y avance en la carrera.
- **Filtros Demográficos:** Segmentos por género, origen, y características sociodemográficas.
- **Interactividad:** Campos clave para visualizaciones y filtros en Power BI.

7. Frontend

- **Interfaz Principal:** Dashboard en Power BI para visualización de datos.
- **Visualizaciones:** Gráficos interactivos (mapas de calor, barras, líneas de tiempo) para análisis de cohortes.
- **Filtros:** año de ingreso, carrera, género y origen para análisis específico.
- **Desglose de Segmentos:** Opciones de desagregación para detalles en cada visualización.
- **Actualización Dinámica:** Datos sincronizados con el DataMart.

8. Autorización de usuario, seguridad de datos

- **Roles de Usuario:** Acceso limitado para analistas y gestores académicos.
- **Autenticación:** Inicio de sesión seguro mediante SSO o autenticación de dos factores.
- **Cifrado de Datos:** Protección de datos sensibles en tránsito y reposo.

- **Cumplimiento Normativo:** Alineación con normativas de privacidad y protección de datos educativos.

9. Soporte de Procesos y Tareas

- **Actualización de Datos:** Procesos automáticos de ETL para cargar datos en el DataMart.
- **Generación de Reportes:** Opciones de exportación y generación de reportes periódicos.

10. Recursos

- **Tecnológicos:** MongoDB para almacenamiento, Power BI para visualización, herramientas de ETL para integración de datos.
- **Humanos:** Analistas de datos, equipo de IT para mantenimiento y soporte.
- **Económicos:** Presupuesto para licencias de software, infraestructura y soporte.

Datos: Acceso a bases de datos de cohortes, rendimiento y características demográficas.

Realizar el proceso de ETL de los datos, para lo cual se debe realizar el proceso de limpieza de los mismos

¿que es el proceso etl?, ¿Cómo se decidieron limpiar los datos?, ¿Por qué?, ¿que decisiones se tomaron?, ¿Cual es el resultado esperado?, ¿Donde se pensaba implementar?, ¿Que metodologia se uso?

Crear un DataMart con los datos asociados en un motor de datos a elección

¿Que es un datam,art?, ¿Como se planifico?, ¿que gestor de base de datos es uso?, ¿Porque?, ¿Hubo problemas en la importacion de los datos y en su implementación?

Desde el DataMart crear un panel de control que se coherente con lapropuesta de valor utilizando PowerBI.

¿que es power bi?, que panel de control se hizo?, porque?, como cumple con la propuesta de valor?, que graficos se usan y porque, ¿Para que categoria de uso fue creado?, (administrador, etc), ¿Porque?

las capturas de pantallas del todo el proceso realizado
Observaciones principales de cada etapa

3. Resultados

¿cuales fueron los resultados del etl?, cuales datos?, cual fue el resultado del data mart?,
¿Cual fue el dashboard resultante?, ¿sirve?, que informacion aporta para tomar decisiones,
porque, que funcionalidades tiene, en que contextos espera ser util y porque, cuales serian
trabajos futuros

Comentarios de los resultados obtenidos en el panel de control

4. Discusión y Conclusiones

Conclusiones personales del equipo respecto al trabajo.

5. Referencias

- [1] CNED, "Índices y estadísticas de la educación superior", [En línea]. Disponible: https://www.cned.cl/indices_New~/pregrado.php?_gl=1*4hps1r*_ga*MTYwNjU0NDMyMi4xNzMwNTc3MDc4*_ga_DJ25RTKCK2*MTczMjA1MDQ3OS42LjEuMTczMjA1MDU1Ni4wLjAuMA [Fecha de acceso: 19 noviembre 2024].