



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

DISEÑO E IMPLEMENTACIÓN DE UNA PROPUESTA ARQUITECTURA DE INGENIERÍA DE DATOS EN LA NUBE PARA LA CONSTRUCCION DE UN REPOSITORIO CENTRALIZADO CON CAPACIDADES (ETL) PARA LA UNIDAD DE EDUCACION PRESENCIAL DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS (ESPE) BASADO EN LA INTEGRACION EFICIENTE DE DATOS EN LA NUBE SIGUIENDO LOS PARAMETROS DE NEGOCIO Y MEJORANDO LA TOMA DE DECISIONES.

AUTOR: CARLOS ANDRÉS RODRÍGUEZ PRECIADO

TUTOR: ING. ANDREA MARGARITA LÓPEZ LÓPEZ, MSc.

SANGOLQUI

2024



1. Introducción
2. Motivación y contexto
3. Diseño Metodológico
4. Evaluación de Resultados
5. Conclusiones y Recomendaciones



1. INTRODUCCIÓN

- 💡 En la actualidad las automatizaciones han sido parte clave de las organizaciones mejorando sus resultados y con la ayuda de arquitecturas en la nube, estos procesos se vuelven más eficientes y seguros.
- 💡 Este estudio propone tener un repositorio centralizado que pueda almacenar información crítica en los procesos de control de asignación de aulas y laboratorios que estos datos son provenientes de los especialistas que manejan en el área de Unidad de Educación Presencial y se busca crear una propuesta de automatización eficiente y segura en la arquitectura de datos.

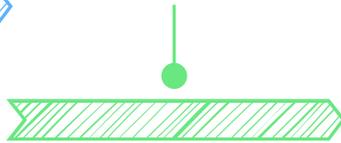


2. MOTIVACIÓN Y CONTEXTO

El manejo de datos en la (UEP) se realiza en hojas de cálculo.



Consecuencias



Tiempo y recursos invertidos

Unión e inconsistencias de datos



Solución



Propuesta de una automatización con esquemas de arquitecturas de datos en la nube sirviendo como un repositorio centralizado y mostrando un análisis en un visualizador.

Beneficios Esperados



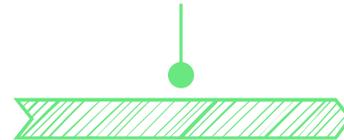
automatización, mejor control, eficiencia y efectividad



Integridad, disponibilidad y precisión de datos



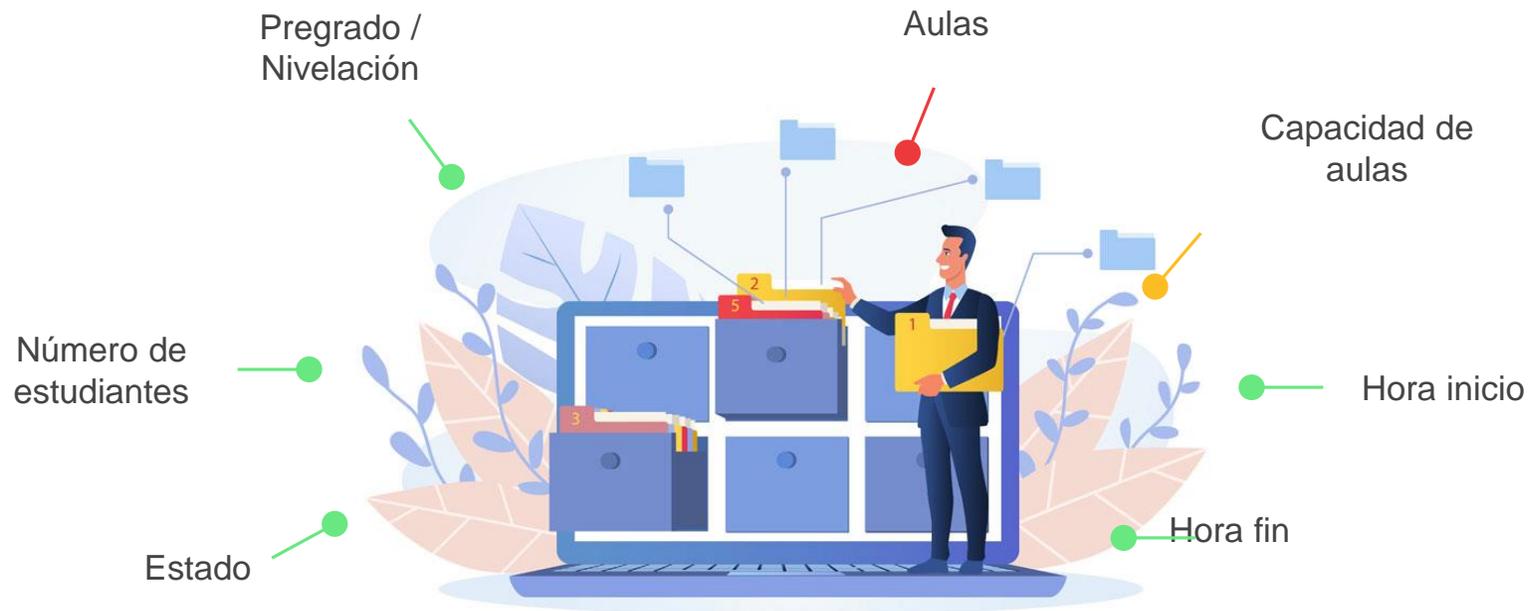
Integración



Integración de todos los datos históricos en un repositorio centralizado



Análisis de Datos



Planificación Por Periodo

Nota. Diccionario y análisis de los datos



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

Objetivo General

- 💡 Automatizar procesos de planificación académica semestral referente a la asignación de infraestructura de la Universidad (aulas y laboratorios) mediante el análisis de datos históricos y proyecciones, creación de un repositorio centralizado y el uso de herramientas de integración de datos en la nube y de visualización con la generación de reportes para la Unidad de Educación Presencial (UEP).



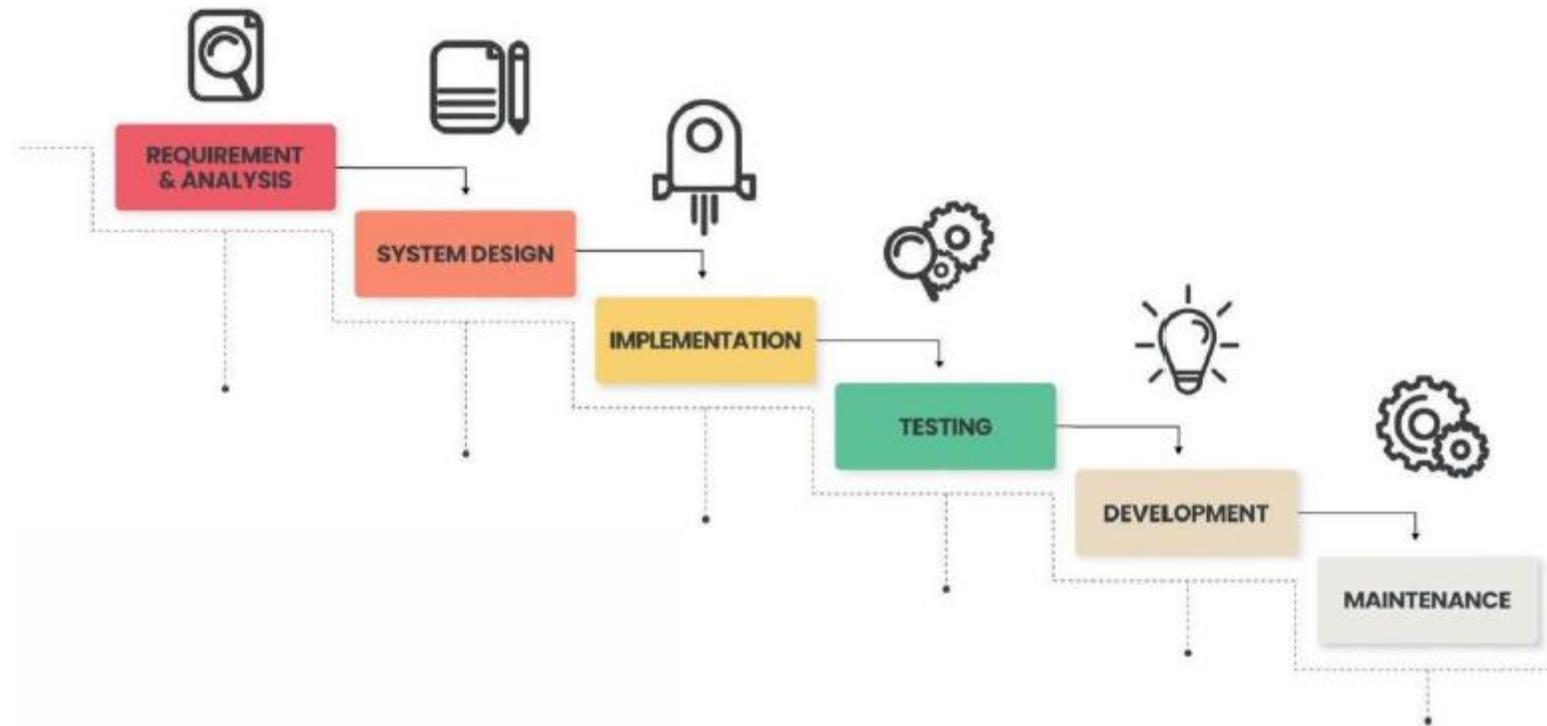
Objetivos Específicos

- 💡 Analizar la data histórica y proyección de datos de planificación académica por semestre sobre procesos de asignación de infraestructura a los departamentos en la Unidad de Educación Presencial (UEP).
- 💡 Diseñar el flujo de datos y los procesos de extracción, transformación y carga (ETL) y su configuración en la nube.
- 💡 Crear y desplegar tableros interactivos en Tableau para la visualización de datos.



3. Diseño Metodológico

Metodología Cascada



Nota. Metodología cascada aplicada en la implementación del proyecto



Herramientas



Grupos de recursos



Máquinas virtuales



Logic Apps



Factorías de datos



Cuentas de almacenamiento



Azure Databricks



Almacenes de claves



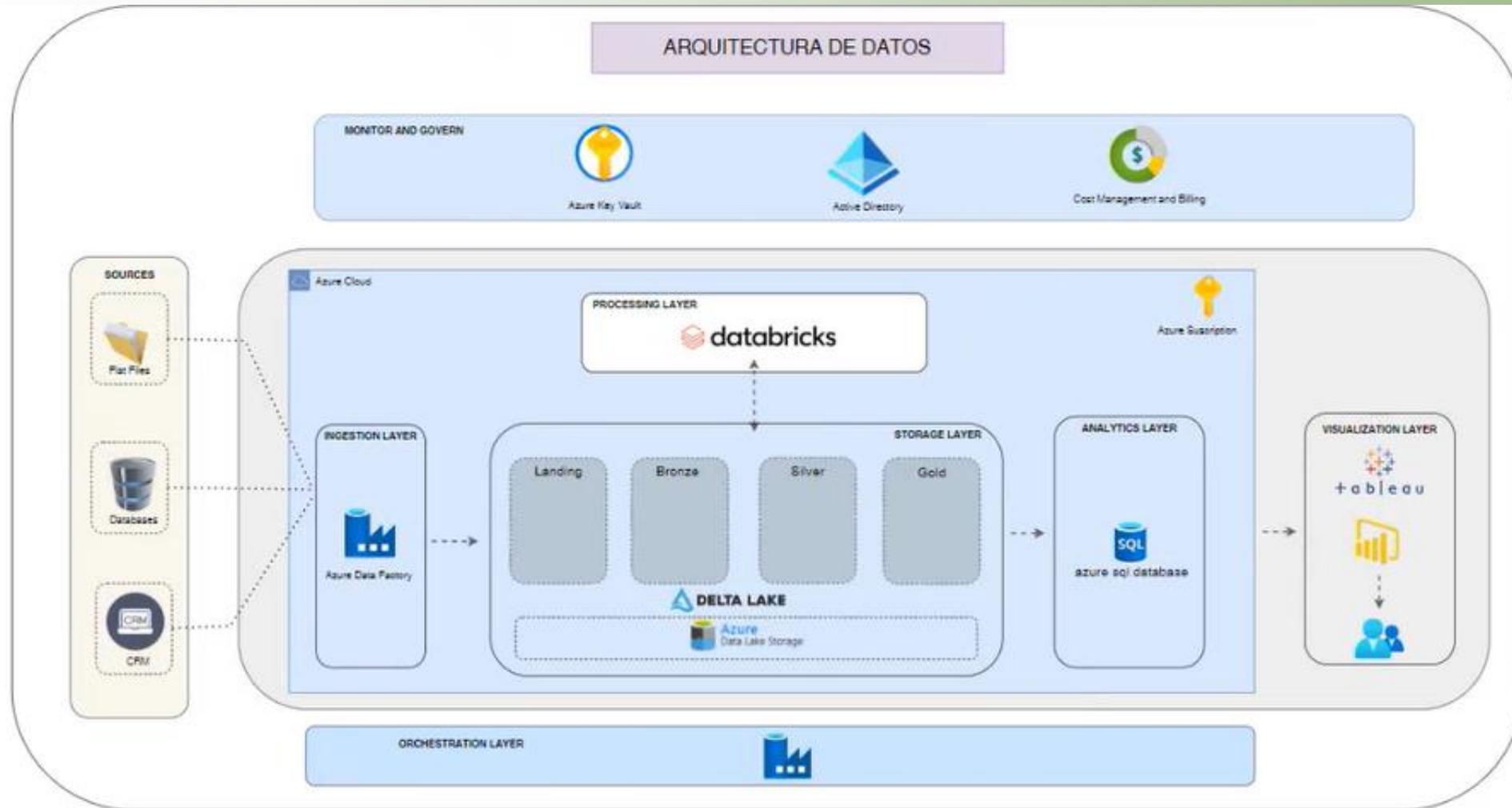
SQL Database



Nota. Descripción de herramientas usadas en el proyecto



ARQUITECTURA DE DATOS



Nota. Arquitectura de datos propuesto



IMPLEMENTACIÓN

Almacenamiento de datos crudos



Azure Data Factory

Máquina Virtual



Integración de on-premise a cloud
cloud a cloud

Contenedor de datos por fases



Data Lake



Azure Databricks

se encargará de al tenerlos datos en la nube limpiarlos, transformarlos, y carga a un base de datos (ETL)

Seguridad y conexión entre servicios



Almacenes de claves



SQL Database

Almacenamiento de datos ya refinados.

Visualizador de resultados



Tableau



Nota. Implementación paso a paso



4. EVALUACIÓN DE RESULTADOS

Resultados

The screenshot displays the Microsoft Fabric Data Factory interface. The main canvas shows a pipeline named 'pipeline2' with two activities: 'EP_EXTRACT PIPELINE_EXTRACT_0...' and 'MOUNT_STORAGE'. A green arrow points from the 'Trigger (1)' icon in the top bar to the 'Edit trigger' dialog box on the right. The 'Edit trigger' dialog shows the following configuration:

- Name: Trigger_ETL_ESPE_002
- Description: (empty)
- Type: ScheduleTrigger
- Start date: 2/16/2024, 10:03:00 AM
- Time zone: Bogota, Lima, Quito (UTC-5)
- Recurrence: Every 1 Week(s)
- Advanced recurrence options: Run on these days (Mon, Tue, Wed, Thu, Fri); Execute at these times (Hours: 8, Minutes: 30); Schedule execution times (08:30)

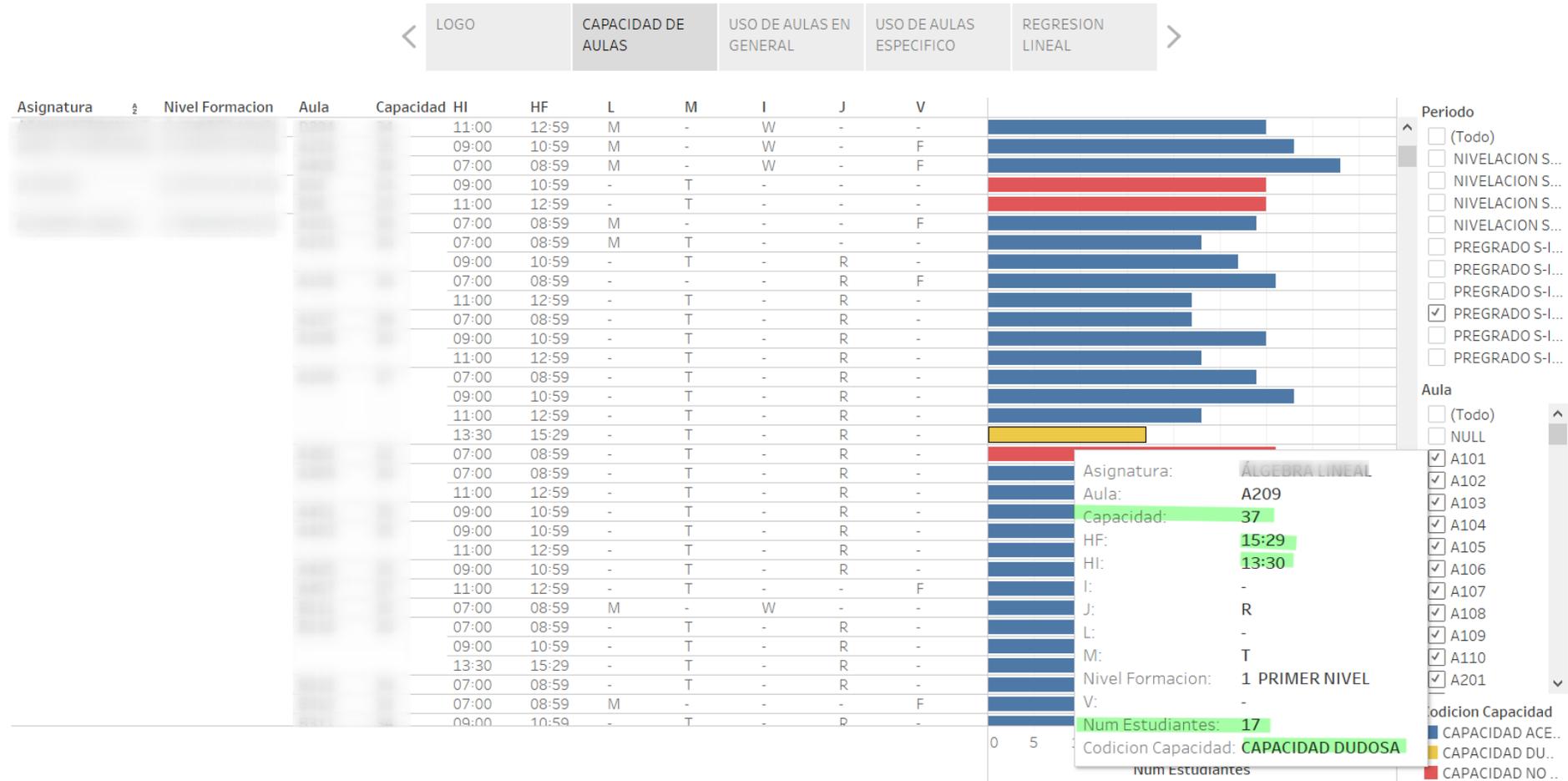
The bottom panel shows the 'Output' tab for the pipeline run. The Pipeline run ID is c1998121-54d5-4bf3-875a-a9fd58824f7. The status is 'All status'. The output table shows the following results:

Activity name	Activity status	Activity type
MAIN_ETL	✔ Succeeded	Notebook
EP_EXTRACT	✔ Succeeded	Execute Pipeline

Nota. Arquitectura levantada y automatizada



Análisis de las Capacidades de las Aulas: hora y día



Nota. Análisis de la capacidad uso en las aulas y laboratorios



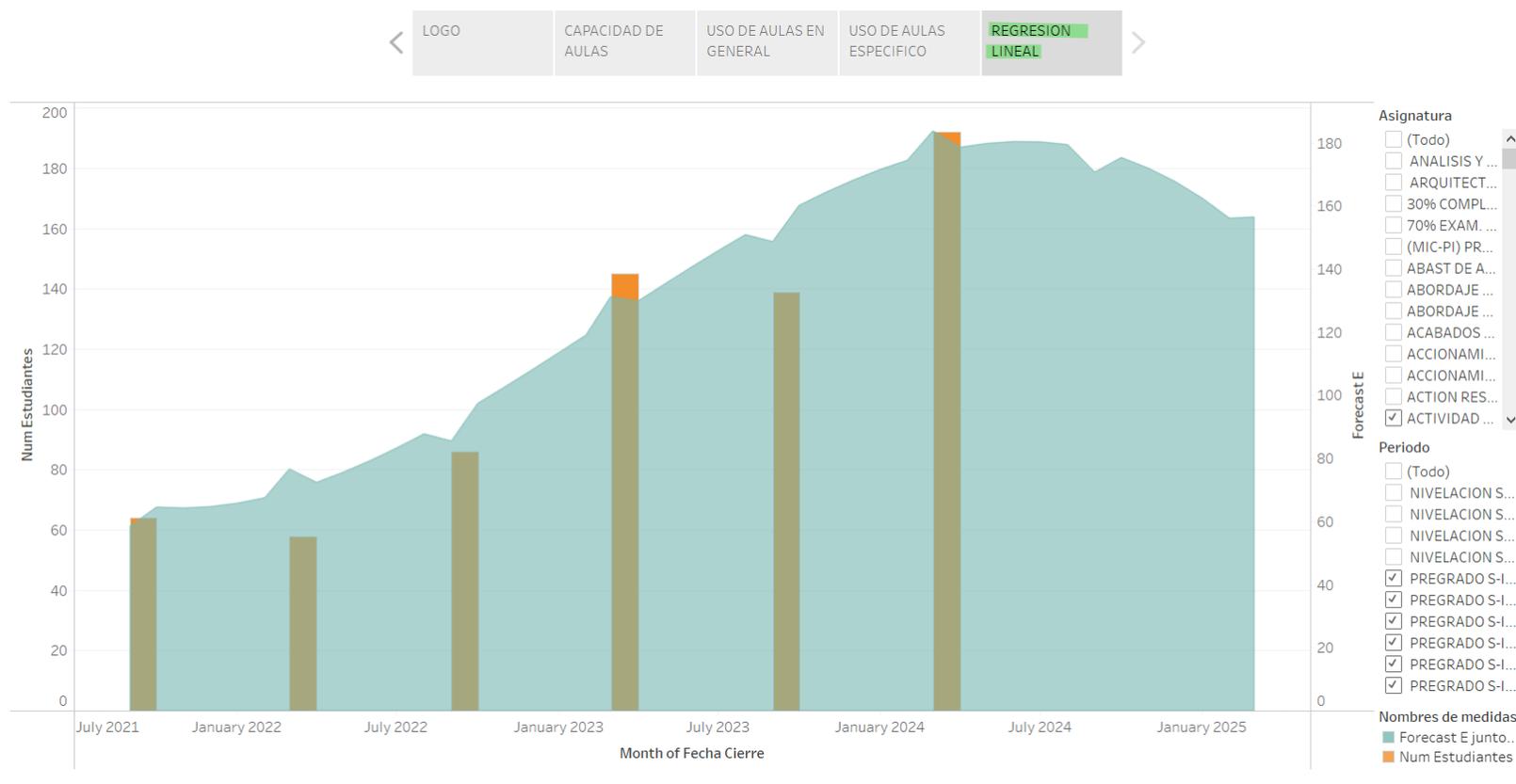
Detalle del uso de Aulas: general y específico por periodo, día y hora



Nota. Análisis del uso de aulas en la universidad con respecto al día y hora



Modelo de Machine Learning con uso de Regresión Lineal de Estudiantes por Materia: Indicador histórico y proyección de estudiantes por materias



Nota. Análisis de regresión lineal sirviendo como proyecciones de la cantidad de alumnos en materias en un futuro



5. CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

- 01 Automatización de procesos → Factible y beneficioso.
- 02 Análisis de data histórica y diseño de flujo de datos → (ETL) en la nube.
- 03 Creación de tableros interactivos con Tableau → Efectiva y Accesible.
- 04 Implementación exitosa → Innovación y eficiencia.



RECOMENDACIONES

Evaluación continua y
Monitoreo de la arquitectura

Tener personal capacitado para el
manejo en la arquitectura de datos

Explorar nuevos procesos
para la automatización

Ampliar un nuevo alcance
de visualización





ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA



GRACIAS POR SU ATENCIÓN

