

A hand in a white sleeve points towards a futuristic digital interface. The background is a dark blue gradient with glowing white gears of various sizes and patterns of light lines, suggesting a complex data analysis process. The overall aesthetic is high-tech and professional.

Proceso de análisis de datos

A decorative graphic in the bottom left corner consisting of several white-outlined hexagons of varying sizes and orientations, some overlapping, creating a molecular or network-like structure.

03 de noviembre de 2020

Formación y experiencia



Nicolás Balparda
INSTRUCTOR

- *Estrategia tecnológica*
- *Arquitectura de datos y sistemas*
- *Big Data*
- *Análisis de datos*

FORMACIÓN

- Ingeniero en sistemas
- MsC en Gestión de Datos e Innovación tecnológica
- MsC en Big Data y Business Analytics (en curso)

Presentación

- a. Nombre
- b. ¿Qué rol ocupa en su trabajo?
- c. ¿Cuál cree que es su relación en el trabajo con los datos?
- d. ¿Participaste en el curso de la semana pasada?

Temario del curso

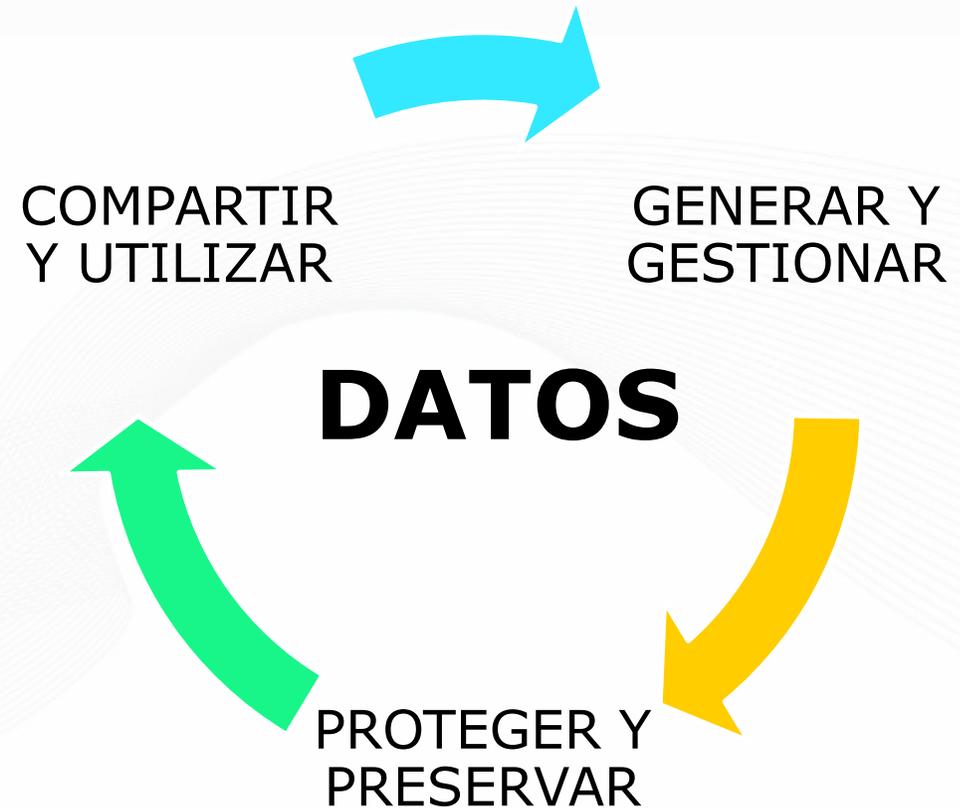
1	Conocimientos previos
	Repaso
2	Proceso de análisis de datos
	Introducción Principales marcos de referencia Descripción de alto nivel
3	Etapas del proceso
	Necesidad Búsqueda de datos Preparación de los datos Modelado y Análisis Presentación Resultados y acciones
4	Roles y Responsabilidades
	Nuevos roles Roles existentes
5	Ejercicio práctico en Excel
	Transformación de datos usando funciones condicionales Analizar los datos utilizando tablas dinámicas y funciones estadísticas Diseño de gráficos
6	Bibliografía y sitios relacionados

Conocimientos previos

Repaso

Ciclo de vida de los datos

La Política de Datos para la Transformación digital fue aprobada en el año 2019 con el objetivo de impulsar una estrategia nacional de datos que promueva y desarrolle proyectos específicos para la gestión de datos en el Estado





¡Repasemos
conceptos!

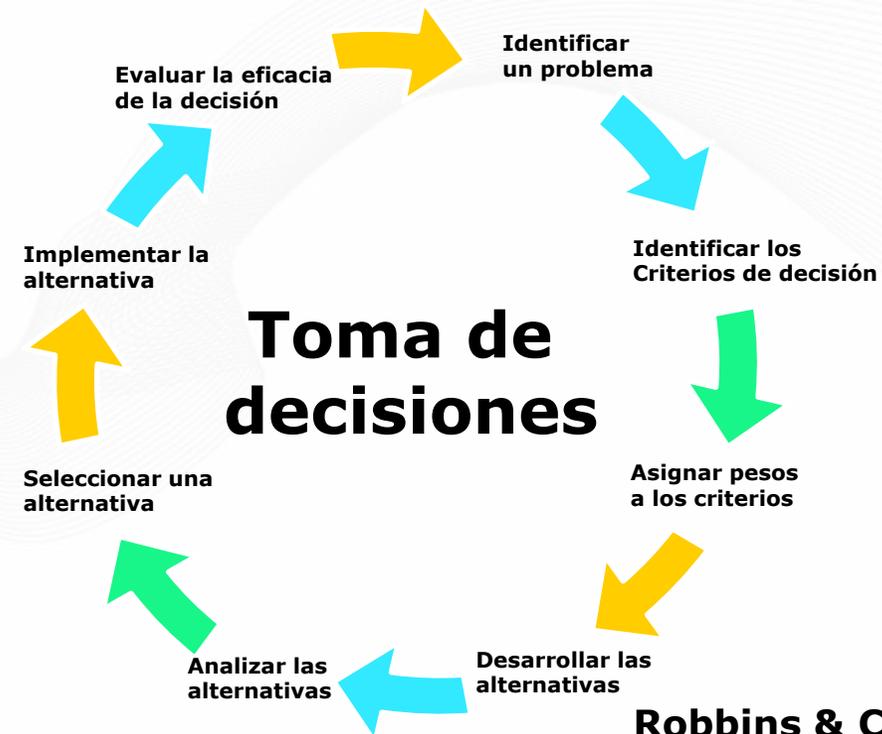
Curso “La importancia de los Datos”:

- **DW:** colección de **datos** bien conocidos y estructurados.
- **ETL:** **extracción, transformación y carga** de datos.
- **BI:** incluye aplicaciones, infraestructura, herramientas y mejores prácticas que permiten el acceso y **análisis de la información** para mejorar y **optimizar las decisiones** y el rendimiento.
- **Big Data:** **activos** de información de **gran volumen, velocidad y / o variedad** que exigen formas rentables e innovadoras de procesamiento de información que permitan una mejor comprensión, toma de decisiones y automatización de procesos.
- **Machine Learning:** mediante la **aplicación de algoritmos** sobre los datos, se permite predecir y encontrar patrones en los mismos.



Proceso de toma de decisiones

La **toma de decisiones** no es solamente la elección entre distintas alternativas, sino que es un **proceso** que incluye distintas etapas:



Robbins & Coulter (2010)

Proceso de análisis de datos



Dinámica

Consigna

- ¿Por qué es importante un proceso de análisis de datos?
- ¿Conocen alguna técnica para hacer un análisis de datos?

15 minutos

Dinámica

¿Por qué el análisis de datos?

La cadena estadounidense Wal-Mart, a principios de los años 90, analizando sus datos con técnicas de DM descubrió que existía un grupo de clientes que, a partir de las 7 de la noche, compraban pañales junto con cerveza. Curiosamente, el perfil del consumidor eran hombres casados entre 25 y 35 años.

Wal-Mart optó por una optimización de las actividades junto a los estantes en los puntos de ventas colocando los pañales al lado de las cervezas.

Resultado: El consumo de cerveza creció 30% con ese cambio. Colocaron snacks y patatas fritas en medio y las ventas de los 3 productos se incrementaron.



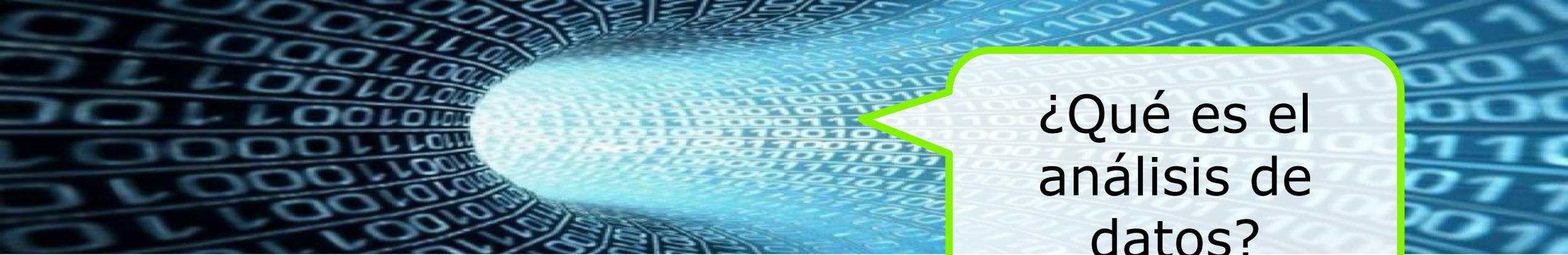
Proceso de análisis de datos

Introducción



Orden vs Caos

Formalizar un **proceso** para maximizar la eficiencia!



¿Qué es el análisis de datos?

El **análisis de datos** es un **proceso** en el que se aplican distintas **técnicas** para **describir, visualizar y evaluar datos**.

El análisis de datos de **Gobierno**:

- Colabora de forma directa e indirecta con una **gestión** gubernamental **eficiente**.
- Políticas **basadas en evidencia**.

¿Por qué un Framework de análisis de datos?

Uruguay

- Plan de Gobierno Digital 2020.
- Política de Datos para la transformación digital.
- AGESIC
 - Arq de Gobierno.
 - Arq de Datos.
 - Plataforma de Datos.
 - Cert.uy

Estándares y Marcos

- Data Management Body of Knowledge (DAMA).
- Open Web Application Security Project (OWASP).
- Marco de Ciberseguridad de AGESIC.

Otros

- Casos de estudio:
 - Regionales.
 - Internacionales.
- Grupos de trabajo e índices:
 - Digital Nations (DN).
 - E-Government Development Index (EDGI)
 - Índice global de ciberseguridad.

Framework de análisis de datos



Objetivo

Brindar lineamientos y recomendaciones prácticas que guíen en el análisis de datos a los organismos estatales.



Gobernanza

- Tipos de análisis.
- Proceso de análisis de referencia.
- Roles y funciones asociadas.
- Mejores prácticas.



Tecnología

- Arquitecturas asociadas a los tipos de análisis.
- Mapeo de las arquitecturas al proceso.
- Mejores prácticas.



Ciberseguridad

- Tipos de análisis.
- Proceso de análisis de referencia.
- Roles y funciones asociadas.
- Mejores prácticas.

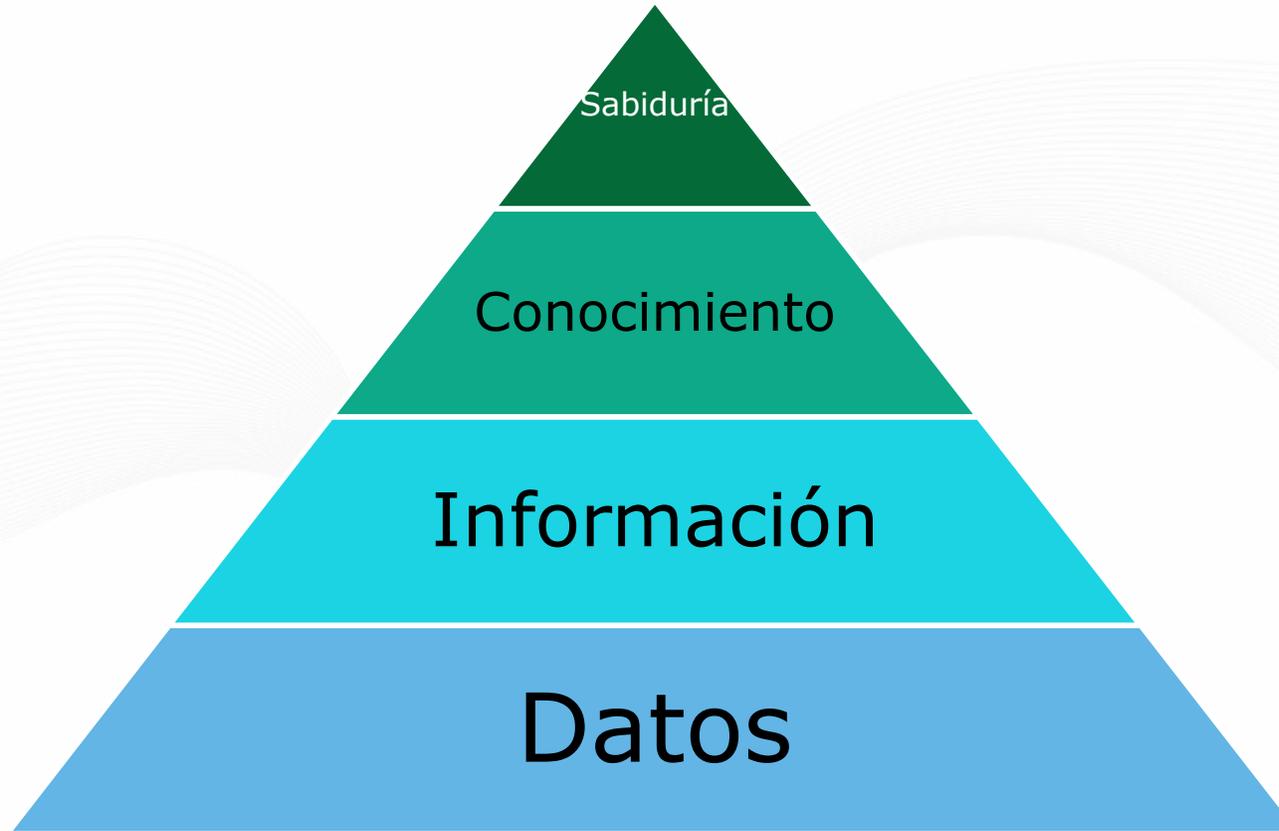


Legal

- Marco legal vigente en Uruguay.
- Consideraciones y recomendaciones para cumplir la normativa sobre los datos personales.

Dimensiones

Tipos de información





Proceso de análisis de datos

Principales marcos de referencia

Marcos de referencia para el proceso

- Procesos para **Data Mining**:
 - Proceso KDD (Knowledge Discovery in Databases).
 - Proceso SEMMA (Sample , Explore, Modify, Model, Assess).
 - Proceso CRISP-DM (Cross-Industry Standard process for Data Mining).
- **Ciencia de datos**:
 - Metodología IBM para ciencia de datos.
- **AGESIC**:
 - Framework de Calidad de Datos.
 - Estrategia de Datos.
 - Estrategia de IA.



Proceso de análisis de datos

The background of the slide features a series of overlapping, wavy lines in a light gray color. These lines create a sense of motion and depth, flowing across the page from left to right. The lines are thin and closely spaced, creating a textured, almost fabric-like appearance.

¿Cómo es el proceso?

¡Veamos **qué hacer** paso a paso!



Data



Analysis



Idea

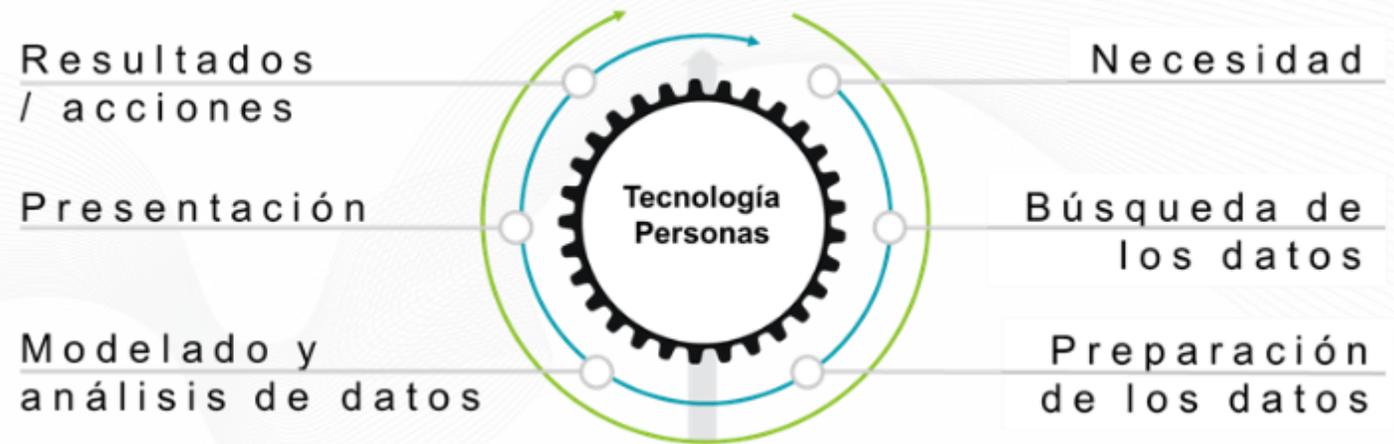


Proceso de análisis de datos

Descripción de alto nivel

Diagrama del proceso de análisis de datos

- Formado por **etapas** que tienen:
 - Objetivos.
 - Entrada.
 - Salida.
 - Actividades.



Etapas del proceso

Objetivos, entrada, salida y actividades

Necesidad

Entrada

Se tiene la necesidad de dar respuesta a una pregunta.

Salida

Documentados los hallazgos de la etapa (necesidad, objetivos, requisitos de la solución, etc.).
Identificados los tipos de análisis y las herramientas a utilizar.

Objetivos

- Identificar las necesidades de información del organismo.
- Guiar las actividades de recolección y posterior análisis.

Buenas prácticas

- Definir criterios de éxito.
- Elección de arquitectura.
- Evaluar procedencia de fuentes de datos.
- Clasificar el tipo de datos.

Actividades

- Definir la necesidad de información.
- ¿Se puede cubrir la necesidad con un análisis pre-existente?
- ¿Existe un antecedente en el que basarse?
- Definir el objetivo, requisitos, criterios de éxito, requisitos de contenido, formato, representaciones, etc.
- ¿El organismo cuenta con las herramientas necesarias para realizar el análisis?

Búsqueda de los datos

Entrada

Documento con las definiciones de la etapa de Necesidad.

Salida

Fuentes de datos recolectadas (en sus fuentes o en una intermedia).

Objetivos

- Definir las fuentes de datos (fuentes confiables).
- Recopilación datos que den respuesta a la necesidad.

Buenas prácticas

- Las fuentes deben estar validadas.
- Seleccionar los datos correctos y pensar datos necesarios a futuro.
- Calidad de las ingestas.

Actividades

- Evaluar si es viable la realización del análisis.
- Seleccionar fuentes internas y/o externas y evaluar su calidad.
- Recolectar los datos.
- Analizar aspectos de los datos (ej.: origen, formato, etc.).

Preparación de los datos

Entrada

Datos recopilados (en sus fuentes o en una intermedia).

Salida

Datos en formato correcto y con niveles de calidad, ingestados en repositorio.

Objetivos

- Procesar, organizar y limpiar los datos para su posterior análisis, utilizando técnicas de transformación, normalización y limpieza de datos.

Buenas prácticas

- Implementar un plan de calidad de datos.
- Identificar outliers.
- Identificar los valores duplicados.
- Enmascarar los datos.

Actividades

- Comprender y explorar los datos mediante técnicas de estadística descriptiva y ciencia de datos (ej.: análisis de clúster y matrices de correlación, etc.).
- Limpieza de los datos (ej.: coincidencia de registros, etc.).
- Normalizar y tratamiento de datos faltantes (ej.: detección de outliers, etc.).
- Ingestar los datos en un repositorio.
- Feature engineering.

Dinámica

Consigna

- Preparación de los datos
 - Tratando datos vacíos o nulos

Dinámica

Preparación de los datos

Tratando datos vacíos o nulos

Cédula	Nombre	Apellido
3.655.789-3	Nicolás	Balparda
3.655.789-4	Nicolás	Balparda
4.111.556-9		Balparda
	Gustavo	Michelakos
	Pablo	Rodríguez
54.893-3	Josefina	Baliño

Preparación de los datos

Tratando datos vacíos o nulos

1. Eliminar los registros que no sean completos.
2. Substituir los valores que no tenemos por otros.
 - a) Reemplazar los valores vacíos con algún valor constante especificado por un analista.
 - b) Reemplazar los valores vacíos por la media (para valores numéricos)
 - c) Reemplazar los valores vacíos con algún valor aleatorio generado a partir de la distribución de la variable.
 - d) Reemplazar los valores vacíos por valores imputados en función de los valores de los otros campos del registro.
3. Trabajar en la fase de modelado con técnicas que sepan tratar de forma robusta el caso de los valores vacíos o nulos.

Preparación de los datos

Identificando clasificaciones incorrectas

Marca	Cantidad
SEAT	1048
CITROEN	658
RENAULT	905
S.E.A.T.	6
RENAUL	3
RENO	1

Preparación de los datos

Identificando clasificaciones incorrectas

Marca	Cantidad
SEAT	1054
CITROEN	658
RENAULT	909

Preparación de los datos

Variables que no aportan nada

Id	País	Edad	Nombre
1	Uruguay	34	Nicolás
2	Uruguay	64	Gustavo
3	Uruguay	63	Graciela
4	Uruguay	33	María
5	Uruguay	33	Ximena
6	Uruguay	48	Roberto

Preparación de los datos

Variables que no aportan nada

Edad	Nombre
34	Nicolás
64	Gustavo
63	Graciela
33	María
33	Ximena
48	Roberto

Preparación de los datos

Transformando variables

Sexo	Nombre
1	Nicolás
-	Gustavo
F	Graciela
M	María
2	Ximena
NULL	Roberto

Preparación de los datos

Transformando variables

Sexo	Nombre	Sexo2
1	Nicolás	Masculino
-	Gustavo	Otro
F	Graciela	Femenino
M	María	Masculino
2	Ximena	Femenino
NULL	Roberto	Otro

Preparación de los datos

Transformando variables

Ingreso	Nombre
USD 9.000	Jaime
\$U 560.000	Esteban
\$U 100.000	Natalia
\$U 90.000	Clara
\$U 30.0000	Julia
\$U 1.000	Daniel

Preparación de los datos

Transformando variables

Ingreso	Nombre	Ingreso2
USD 9.000	Jaime	\$U 387.000
\$U 560.000	Esteban	\$U 560.000
\$U 100.000	Natalia	\$U 100.000
\$U 90.000	Clara	\$U 90.000
\$U 30.0000	Julia	\$U 30.000
\$U 1.000	Daniel	\$U 1.000

Preparación de los datos

Transformando variables

Ingreso	Nombre	Ingreso2
USD 9.000	Jaime	\$U 387.000
\$U 560.000	Esteban	\$U 560.000
\$U 100.000	Natalia	\$U 100.000
\$U 90.000	Clara	\$U 90.000
\$U 30.0000	Julia	\$U 30.000
\$U 1.000	Daniel	\$U 1.000

Preparación de los datos

Transformando variables

Supongamos un proceso industrial que se ve afectado ligeramente por la temperatura ambiental. Trabajar con el valor de la temperatura exacto posiblemente es innecesario y añade poco valor. Sería más adecuado categorizar esa variable en una que indique un rango de temperaturas. ¿Cómo lo harían?

Temp
- 10 °C
- 10,5 °C
10 °C
16 °C
18 °C
26 °C

Preparación de los datos

Transformando variables

Temp	Temp 2
- 10 °C	Muy frío
- 10,5 °C	Muy frío
10 °C	Frío
16 °C	Templado
18 °C	Templado
26 °C	Caluroso

Modelado y análisis de datos

Entrada

Datos preparados en un almacén de datos.

Salida

Documentación del modelo y de los resultados.

Objetivos

- Definir y construir modelos de análisis.
- Analizar los datos para obtener los objetivos de la etapa de Necesidad.

Buenas prácticas

- Utilizar indicadores estándar de la industria.
- Entrenamiento de modelos.
- Monitorear los modelos.
- Política de revisión de datos.

Actividades

- Técnicas de análisis de datos (ej.: modelos multidimensionales, algoritmos, etc.).
- Desarrollar hipótesis y métodos de testeo.
- Utilizar modelos para encontrar tendencias y correlaciones (dependen de los datos y del enfoque elegido).
- Evaluar si el modelo responde la problemática (funcional, cuantitativa).

Presentación

Entrada

Información resultante del análisis y los requerimientos de la etapa de Necesidad.

Salida

Los resultados están disponibles.

Objetivos

- Comunicar los resultados el análisis en el formato acordado.

Buenas prácticas

- Para la visualización usar colores y fuentes consistentes.
- Conocer a la audiencia.
- Capacitar a usuarios finales en las herramientas de visualización.

Actividades

- Definir qué visualización es adecuada para el público objetivo.
- Diseñar las visualizaciones utilizando herramientas.
- Disponibilizar la información.

Resultados y acciones

Entrada

Documentación sobre el análisis, visualización y definiciones realizadas a lo largo de proceso.

Salida

Acciones, definiciones y decisiones.

Objetivos

- Evaluar los resultados.
- Tomar decisiones.

Buenas prácticas

- Informar sobre la disponibilidad del análisis.
- Gestionar el conocimiento.
- Establecer procesos para recibir, analizar y responder los reportes.

Actividades

- Verificar si la información recolectada es útil para satisfacer la Necesidad.
 - ¿El análisis responde las preguntas iniciales?
 - ¿Es concluyente?
 - ¿Surgieron nuevas necesidades?
- Evaluar resultados y tomar decisiones.

Roles y Responsabilidades



A low-angle, top-down photograph of a diverse group of people of various ages and ethnicities. They are all smiling and looking towards the center, where their hands are stacked together in a circle. The lighting is bright and warm, creating a positive and collaborative atmosphere. The background is a plain, light-colored wall.

Hay equipo!

El equipo de trabajo
es **fundamental**
¿Cómo es tu
equipo?



Nuevos roles

Científico de datos.
Arquitecto de análisis de datos.
Ingeniero de datos.
Analista de información.
Custodio del análisis de datos.

Roles existentes

Responsable de Seguridad de la Información.
Delegado de protección de datos.
Tomador de decisiones.
Encargado de dominio.

Roles y Responsabilidades

Nuevos roles

Científico de datos



Misión

Extraer datos, internos y externos, para analizarlos con el fin de generar valor a la organización.

Formación

- Ciencias económicas, ingeniería o afines.
- Estadística / matemática.

Competencias

- Programación.
- Funciones del organismo.
- Herramientas de análisis.
- Estadística.
- Machine Learning.
- Modelos predictivos.

Actividades

- Recibir requerimientos por parte de los interesados y del Analista de información.
- Sugerir herramientas, modelos de análisis y datos.
- Aplicar técnicas de análisis.

Arquitecto de análisis de datos



Misión

Conoce e identifica herramientas de análisis de datos a incorporar en el organismo, con el fin de optimizar costos y esfuerzos en el uso de los recursos.

Formación

- Analista en computación.

Competencias

- Base de datos relacionales y no relacionales.
- Arquitectura de Software.
- Funciones del organismo.
- Soluciones de data warehousing.
- Manejo de activos de gobierno digital.
- Arquitectura del organismo y de la Arquitectura Integrada de gobierno.

Actividades

- Diseñar la arquitectura de análisis de datos.
- Evalúa herramientas de análisis y sistemas externos.

Ingeniero de datos



Misión

Diseñar, desarrollar, testear y mantener sistemas de análisis de datos.

Formación

- Programador.

Competencias

- ETL.
- Scripting.
- Programación.
- Base de datos relacionales y no relacionales.
- Soluciones de data warehousing.

Actividades

- Programar ETLs.
- Investigar herramientas de ETL.
- Diseñar sistemas de información junto con el arquitecto de análisis de datos.
- Gestionar el acceso a las fuentes.
- Gestionar la calidad de los datos.

Analista de información



Misión

Crear dashboards, tableros, reportes y modelos de análisis según los requerimientos obtenidos de los interesados.

Formación

- Ciencias económicas, ingeniería o afines.

Competencias

- Funciones del organismo.
- Programación básica.
- Herramientas de análisis.
- Consultas SQL.

Actividades

- Trabajar con los interesados.
- Investigar las funciones del organismo para prever qué se puede necesitar.
- Diseñar visualizaciones estáticas y/o interactivas.
- Explorar necesidades y resolverlas mediante el análisis de datos.
- Nexo entre el Ingeniero de datos y los interesados.
- Gestionar impactos en los análisis actuales ante posibles cambios.

Custodio del análisis de datos



Misión

Responsable de la gestión de los datos. Manejar procesos de datos, política de datos y realizar guías sobre el uso de los mismos.

Formación

- N/A

Competencias

- Manejo de bases de datos relacionales y no relacionales.
- Funciones del organismo.
- Arquitectura de datos.
- Manejo de activos de gobierno digital.
- Procesos de calidad de datos.

Actividades

- Disponibilidad de los datos.
- Seguridad de los datos.
- Optimizar el almacenamiento.
- Diseñar, gestionar y controlar procesos.
- Establecer métricas y requisitos de calidad. Velar por el cumplimiento en todas las etapas del proceso.

Roles y Responsabilidades

Roles existentes

Responsable de Seguridad de la Información



Misión

Gestiona, de forma integrada, las medidas de ciberseguridad con perspectiva de riesgo en la organización.

Formación

- Ingeniero / Licenciado en computación.

Competencias

- Auditoría y gestión de riesgo.
- Gestión de incidentes.
- Redes y Hardware.
- Arquitectura de datos.
- Arquitectura del organismo y la integrada de gobierno.

Actividades

- Prevención, detección y control de riesgo.
- Gestiona de forma proactiva y reactiva los riesgos.
- Diseñar planes de recuperación y contingencia.
- Investiga brechas de seguridad.
- Planifica, supervisa y gestiona hackeos éticos.
- Gestión de acceso.
- Implementar técnicas de autenticación y autorización.

Delegado de protección de datos



Misión

Validar que el análisis de datos respete el marco regulatorio y legal vigente

Competencias

- Marco regulatorio y legal sobre el análisis de datos.
- Competencias del organismo.

Actividades

- Monitorear que se cumplen los reglamentos y los estándares.
- Resolver cuestiones legales referidos al análisis de datos.

Formación

- Abogado.

Tomador de decisiones



Misión

Revisar reportes, tableros e indicadores para validar su correctitud. Tomar decisiones a partir del análisis.

Formación

- N/A

Competencias

- Conocimientos de los datos disponibles en la organización o fuera de ella.
- Interpretación de datos.

Actividades

- Revisar KPIs y dashboards para sacar conclusiones y actuar en consecuencia.

Encargado de dominio



Misión

Nexo entre el área usuaria, los interesados y los Analistas de Información.

Formación

- Estudios que competan al organismo.

Competencias

- Experto en las funciones del organismo según su dominio (Marketing, Operaciones, Finanzas).
- Conocer qué datos están disponibles y en qué sistemas.

Actividades

- Recomendaciones sobre variables y formas de presentar los datos.
- Recomendar fuentes para su análisis.
- Evaluar, junto con el Custodio de los datos y el Ingeniero de datos, la calidad de los datos.
- Validar los datos del análisis, verificando su coherencia con la realidad.
- Apoyar en la interpretación de los datos y su análisis.

Ejercicio práctico en Excel

Transformación de datos usando funciones condicionales

Analizar los datos utilizando tablas dinámicas y funciones estadísticas

Diseño de gráficos

1. Necesidad

¡Apliquemos el proceso!

Dada la **encuesta continua de hogares** 2019 del INE queremos responder las siguientes preguntas:

1. ¿Cuál es el **departamento** con el porcentaje más **alto** de personas que terminaron **estudios universitarios**?, ¿y el más bajo?
2. ¿Existe **relación** entre el resultado anterior y los **ingresos**?



2. Búsqueda de los datos

¡Busquemos y entendamos los datos!

Planilla **personas**:

- numero: identificador del hogar.
- nomdpto: departamento de la persona.
- e218_1: 1 si terminó la universidad, 2 sino.
- e220_1: código de la carrera que cursó.

Planilla de **hogares**:

- numero: identificador del hogar.
- YSVL: ingreso total del hogar.

Especificación de **carreras**:

- Nivel + Destino + Sector: código de la carrera.
- Credito educativo: descripción de la carrera.

3. Preparación de los datos

¡Preparemos los datos!

¿**Todos** los **datos** salen de la fuente de **Personas**?

¿Cómo **vinculamos** las **Personas** con los **Hogares**?

- Buscar el atributo que tienen en común.
- Analizar si hay columnas que se repiten.
- Utilizar la función buscarv (vlookup)

=*VLOOKUP(<IDHogar>;<hogares>;<colIngreso>;TRUE)*

¿Cómo **vinculamos** las **Personas** con las **Carreras**?

- Armarnos el atributo por el que las vamos a buscar concatenando Nivel + Destino + Sector (ID)
- Utilizamos la función buscarv (vlookup).

=*VLOOKUP(<ID>;<carreras>;<colDescripcion>;TRUE)*



¡Comparemos los departamentos!

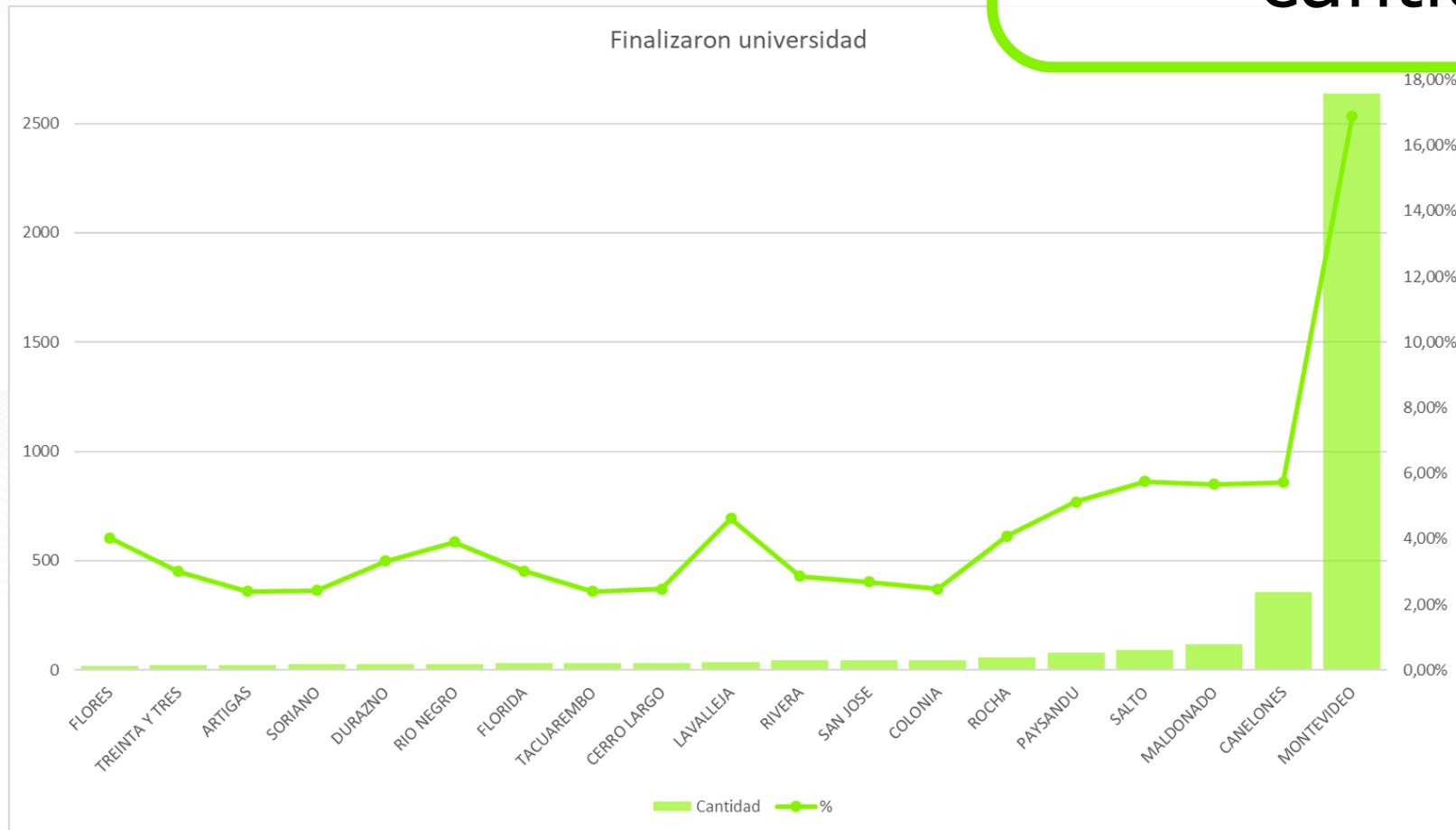
¿Cuál es el **departamento** con el porcentaje más **alto** de personas que terminaron **estudios universitarios**?, ¿y el más bajo?

1. Creemos una tabla dinámica: Insert -> PivotTable.
 - a) Arrastremos el departamento a las filas.
 - b) Arrastremos la variable "Cursó universidad" a Columnas para ver los 3 valores.
 - c) Arrastremos el ID del hogar para que sume la cantidad de hogares por departamento.
 - d) Ídem c, poniendo como porcentaje (relativo).
2. Creemos una escala de valores sobre los que finalizaron la universidad (Data Bars parados en la columna de porcentaje).

Finalizaron universidad		
	Cant	%
Departamento	Terminó	Terminó
ARTIGAS	24	2,40%
CANELONES	358	5,73%
CERRO LARGO	32	2,48%
COLONIA	45	2,47%
DURAZNO	27	3,33%
FLORES	18	4,03%
FLORIDA	30	3,02%
LAVALLEJA	37	4,63%
MALDONADO	117	5,67%
MONTEVIDEO	2639	16,89%
PAYSANDU	81	5,13%
RIO NEGRO	28	3,90%
RIVERA	43	2,87%
ROCHA	56	4,08%
SALTO	93	5,75%
SAN JOSE	43	2,69%
SORIANO	26	2,43%
TACUAREMBO	30	2,40%
TREINTA Y TRES	21	3,00%

Analicemos el gráfico

¿Por qué hay diferencias entre % y cantidad?



4. Modelado y análisis de datos

¡Busquemos relaciones!

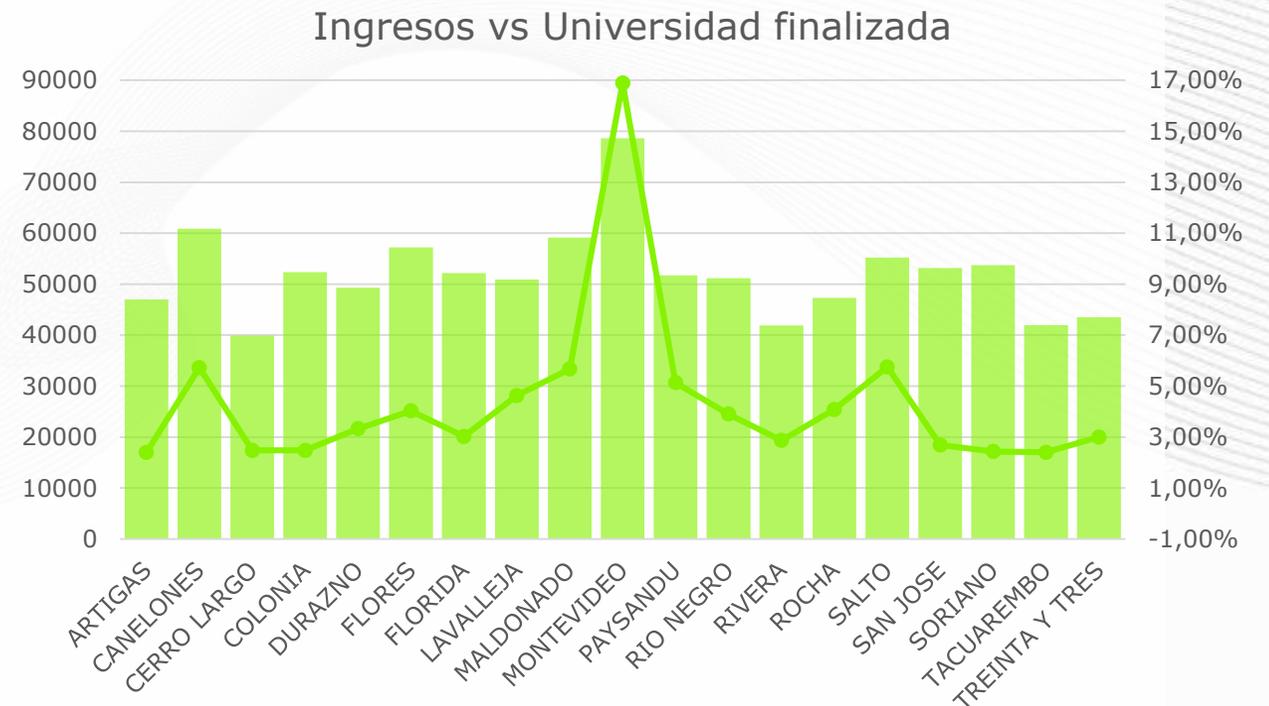
¿Existe **relación** entre el resultado anterior y los **ingresos**?

1. Creemos una tabla dinámica: Insert -> PivotTable.

a) Arrastremos el departamento a las filas.

b) Arrastremos el ingreso y pongámoslo como promedio.

2. Comparemos los ingresos y la finalización de la universidad con una gráfica.



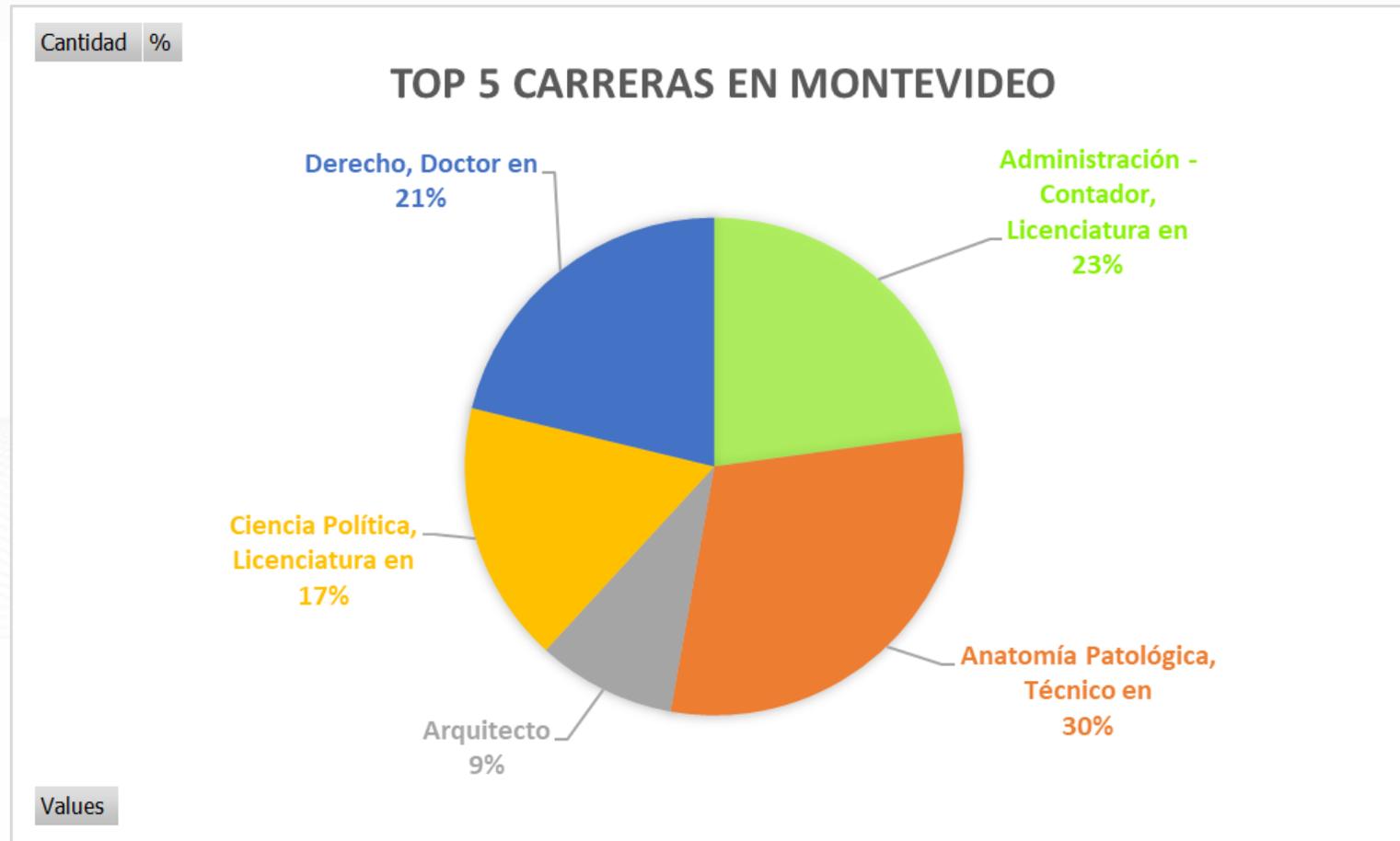
¿Terminamos?

Al presentar los análisis, pueden surgir **nuevas preguntas** y esto nos llevaría a volver a iniciar el ciclo de análisis:

Para el departamento con mayores personas que terminaron la universidad, ¿**cuáles** son las cinco **carreras más elegidas**?

- Creemos una tabla dinámica para visualizar la elección de las carreras por departamento.
- Creemos una tabla dinámica para visualizar la elección de las carreras.
- Linkeamos las tablas a través de los filtros (Insert Slide y luego Report Connections) de Departamento y Cursó universidad. Seleccionemos “Montevideo” y “1”.
- Tomemos el top 5 de las carreras más elegidas.
- Insertemos una gráfica para visualizar la elección de las carreras (se aplican los filtros).

¡Elección de Carreras!



5. Resultados / acciones

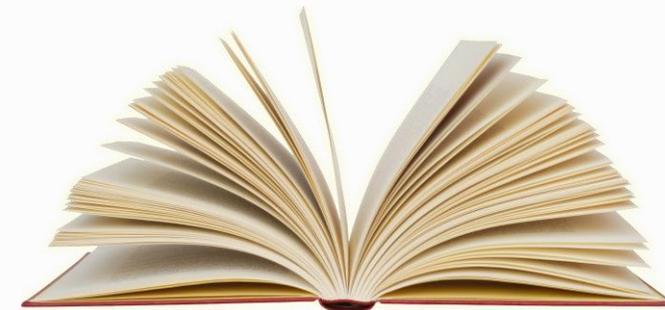
¡Tomemos
decisiones!

Sabiendo que **Artigas** y **Tacuarembó** tienen el **menor porcentaje** de personas que **finalizaron carreras universitarias**, ¿**cómo** podemos **mejorar** sus números?

DEPTOS Y CARRERAS	Cantidad	%
ARTIGAS	24	100,00%
Administración - Contador, Licenciatura en	6	25,00%
Aeronáutica, Técnico de	1	4,17%
Anatomía Patológica, Técnico en	8	33,33%
Ciencia Política, Licenciatura en	1	4,17%
Ciencias de la Educación, Licenciatura en	1	4,17%
Derecho, Doctor en	3	12,50%
Ingeniero Agrónomo	2	8,33%
Oficial de policía	2	8,33%

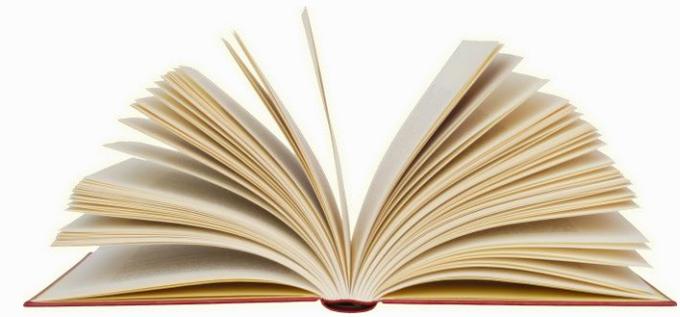
DEPTOS Y CARRERAS	Cantidad	%
TACUAREMBO	30	100,00%
Anatomía Patológica, Técnico en	5	16,67%
Arquitecto	2	6,67%
Ciencias de la Educación, Licenciatura en	2	6,67%
Ciencias Históricas, Licenciatura en	1	3,33%
Ciencias Veterinarias, Doctor en	4	13,33%
Derecho, Doctor en	5	16,67%
Ingeniero Agrónomo	8	26,67%
Oficial de policía	1	3,33%
Servicio Social, Licenciatura en	2	6,67%

Referencias y sitios relacionados



- 1) [Sternkopf H, Mueller R \(2018\). Doing Good with Data: Development of a Maturity Model for Data Literacy in Non-governmental Organizations. *Proceedings of the 51st Hawaii International Conference on System Sciences*.](#)
- 2) [Ridsdale C, Rothwell J, Smith M, Ali-Hassan, Bliemel M, Irvine D, Kelley D, Matwin S, Wuetherick B. Strategies and Best Practices for Data Literacy Education. *Dalhousie University*.](#)
- 3) [Grillenberger A, Romeike R \(2018\). Developing a Theoretically Founded Data Literacy Competency Model. *WiPSCE*.](#)
- 4) [Data to the people \(2018\). Databilities. *Sitio web Wix*.](#)
- 5) [Bonikowska A, Sanmartin C, Frenette M \(2019\). Data Literacy: What It Is and How to Measure It in the Public Service. *Sitio web Statics Canada*.](#)
- 6) [*Sitio web Data Literacy Project*.](#)
- 7) [Department of the Prime Minister and Cabinet \(2016\). Data skills and capability in the Australian Public Service. *Sitio web Australian Government*.](#)
- 8) [*Sitio web Open Data Institute*.](#)
- 9) [AGESIC, Presidencia de la República Oriental del Uruguay \(2019\). Plan de Gobierno Digital 2020: Transformación con equidad. *Sitio web Presidencia de la República Oriental del Uruguay*.](#)

Referencias y sitios relacionados



- 10) DAMA International (2019). DAMA – DMBOK: Data Management Body of Knowledge. *Technics Publications*.
- 11) [AGESIC \(2019\). Uruguay: Política de Datos para la Transformación Digital. Sitio web Presidencia de la República Oriental del Uruguay.](#)
- 12) [Azevedo A, Santos M.F. \(2008\). KDD, SEMMA and CRISP-DM: A PARALLEL OVERVIEW.](#)
- 12) [IBM Analytics. Foundation Methodology for Data Science IBM.](#)
- 13) [AGESIC \(2019\). Framework de Análisis de Datos. Sitio web Presidencia de la República Oriental del Uruguay.](#)
- 14) [AGESIC \(2019\). Marco de referencia para la gestión de calidad de datos. Sitio web Presidencia de la República Oriental del Uruguay.](#)
- 15) [Política y estrategia de datos para la transformación digital \(2018\). AGESIC. Sitio web Presidencia de la República Oriental del Uruguay.](#)
- 16) [AGESIC \(2019\). Estrategia de Inteligencia Artificial para el Gobierno Digital. Sitio web Presidencia de la República Oriental del Uruguay .](#)
- 17) [Robbins S, Coulter M \(2010\). Administración Décima edición. Pearson Educación, 121-124.](#)
- 18) [Rodríguez-Ponce E, Pedraja-Rejas L \(2009\). Análisis del impacto del proceso de toma de decisiones estratégicas sobre la eficacia de las organizaciones públicas. INNOVAR Revista de Ciencias Administrativas y Sociales, 19\(35\), 33-46.](#)
- 19) [Cañón V, Clavijo A, Godoy L, Letouzé E, Pestre G, Ricard J \(2017\). Definición de la estrategia de Big Data para el estado colombiano y para el desarrollo de la industria de big data en Colombia 2017 - 2018. Sitio web Data-Pop Alliance.](#)

Muchas gracias.