



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Introducción a Metodologías



Laboratorio de Datos
2do cuatrimestre 2024



Recorrido de la materia (hasta ahora)

- Lenguaje de programación para trabajar en nuestros proyectos



Actividad N° 1 - Trabajo en equipo

- Conformar grupos de 3 integrantes
- A partir de los datos del archivo Clase-03-Actividad-01-Datos.csv (se encuentra subido al campus) **escriban sus conclusiones.**
- ¡NO SE PUEDEN HACER PREGUNTAS A LXS DOCENTES!
- Tienen 20 minutos

Actividad N° 1 - Puesta en común

1. ¿Qué conclusiones obtuvieron?
2. ¿Con qué problemas se enfrentaron?
3. ¿De qué manera trabajaron? (pasos que siguieron)
4. ¿Todos los integrantes cumplieron el mismo rol?

| | | |
|--|--|--|
| | | |
|--|--|--|

Actividad N° 1 - Reformulando las preguntas

1. ¿Qué estaban buscando? ¿Cuál era la pregunta que intentaban responder?
2. ¿Había alguien en el grupo que conocía del tema? → Dominio del problema
3. ¿Contaban con los datos necesarios para responder la pregunta?
4. ¿Había datos de más?

| | | |
|--|--|--|
| | | |
|--|--|--|

Actividad Nro. 1 - Fuente de datos

Los datos con los que trabajaron fueron obtenidos a partir siguiente link:

[https://datos.gob.ar/dataset/jgm-puntos-wifi-pais-digital/archivo/jgm_13.](https://datos.gob.ar/dataset/jgm-puntos-wifi-pais-digital/archivo/jgm_13.2)

2

The screenshot displays the 'datos.gob.ar' website interface. At the top, the logo 'datos.gob.ar' is on the left, and navigation links 'Datasets', 'Series', 'Organizaciones', 'APIs', and 'Acerca' are on the right. The main content area shows the dataset 'Puntos WiFi (XLSX)' with a description: 'Listado de los puntos de acceso a internet de la Red WiFi de Pais Digital (XLSX)'. A green 'DESCARGAR' button is visible. Below this, the 'Información adicional' section lists details: 'Responsable' (Jefatura de Gabinete de Ministros, Secretaría de Innovación Pública, Subsecretaría de Servicios y Pais Digital), 'Formato' (xlsx), 'Temas' (Ciencia y tecnología, Educación, cultura y deportes, Gobierno y sector público), 'Frecuencia de actualización' (Eventual), 'Fecha de creación' (31 de mayo de 2019), and 'Último cambio' (desconocido). At the bottom, the 'Otros recursos de este dataset' section shows a link to 'Puntos WiFi (CSV)' with a description: 'Listado de los puntos de acceso a internet de la Red WiFi de Pais Digital (CSV)'. A blue 'CONSULTAR' button and a blue 'DESCARGAR' button are also present.

datos.gob.ar Datasets Series Organizaciones APIs Acerca -

Datos Argentina / Dataset / Recurso

Puntos WiFi (XLSX) DESCARGAR

Listado de los puntos de acceso a internet de la Red WiFi de Pais Digital (XLSX).

Información adicional

Responsable Jefatura de Gabinete de Ministros. Secretaría de Innovación Pública. Subsecretaría de Servicios y Pais Digital

Formato **xlsx**

Temas Ciencia y tecnología Educación, cultura y deportes Gobierno y sector público

Frecuencia de actualización Eventual

Fecha de creación 31 de mayo de 2019

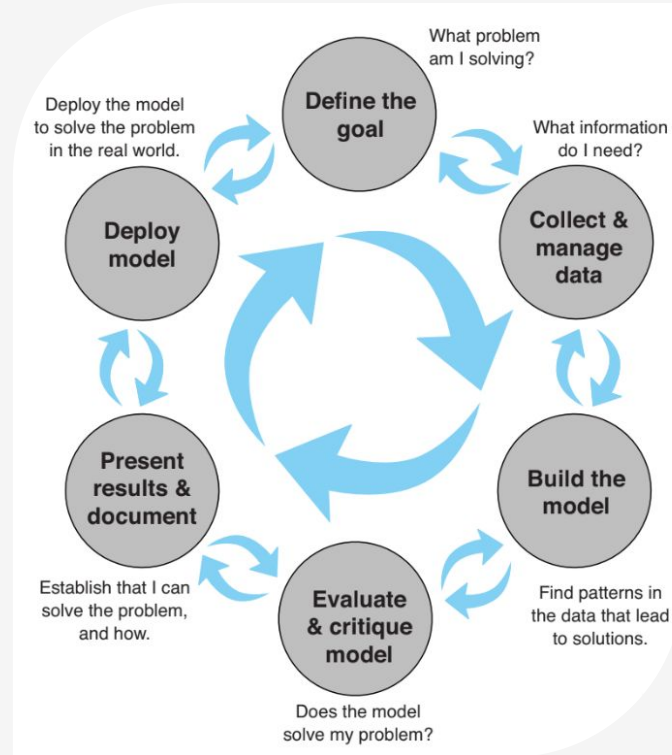
Último cambio desconocido

Otros recursos de este dataset

CSV Puntos WiFi (CSV) CONSULTAR

Listado de los puntos de acceso a internet de la Red WiFi de Pais Digital (CSV). DESCARGAR

Posibles etapas de un proyecto de Ciencia de Datos

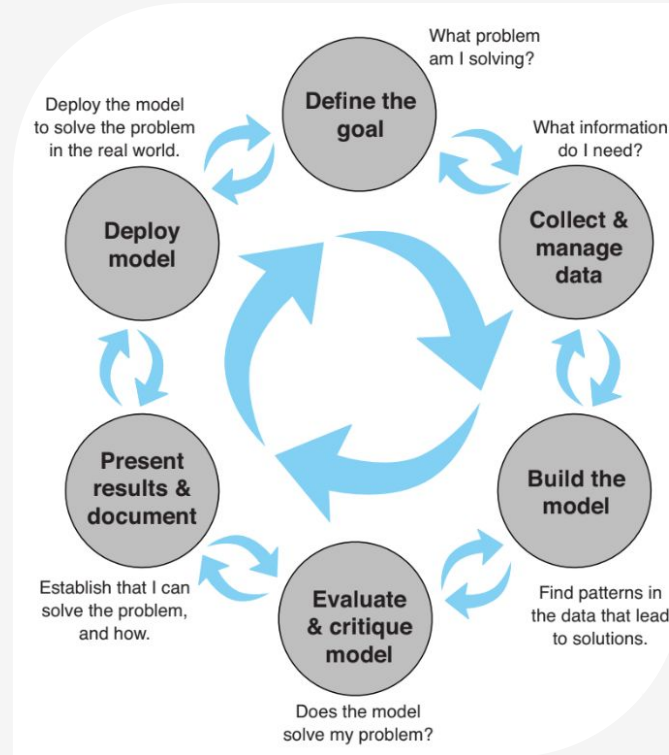


Actividad Nro. 1 - Reflexiones

- Seguir una metodología ordena y permite evaluar si se cumplió con el objetivo
- Existen distintas metodologías:
 - CRISP-DM (Cross Industry Standard Process for Data Mining)
 - KDD (Knowledge Discover Database)
 - SEMMA (Sample, Explore, Modify, Model and Assess)

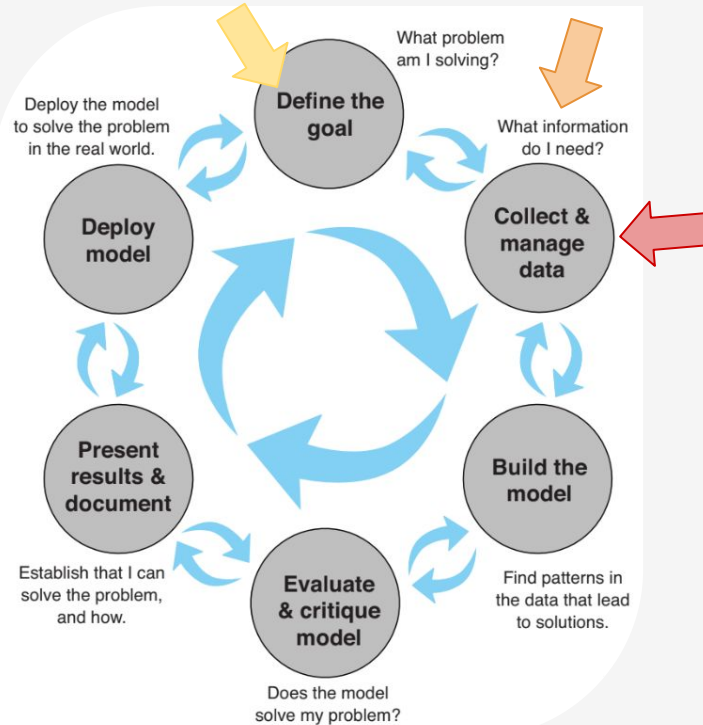
| | | |
|--|--|--|
| | | |
|--|--|--|

Posibles etapas de un proyecto de Ciencia de Datos



1er parte de la materia

Actividad Nro. 2 - Consigna



1. Conformar grupos de 3 integrantes
2. Objetivo: Conocer cuál es el transporte más usado (por los alumnos de nuestro curso) para llegar a Ciudad Universitaria
3. Decidir qué datos debemos recopilar
4. Pensar cómo a partir de los datos a recopilar cumpliremos el objetivo (proceso a realizar)

Tienen 20 minutos

Actividad Nro. 2 - Debate

1. ¿Qué datos decidieron recopilar?
2. ¿Qué pasa con los casos especiales? Por ejemplo, qué pasa con los que utilizan más de un transporte? ¿Consideramos sólo el tramo final? ¿Y los que utilizan un transporte distinto según el día?
3. ¿Qué herramienta utilizarán para recopilar los datos?
4. ¿De qué manera esperan almacenar dicha información?
5. ¿Cómo la procesarán?

| | | |
|--|--|--|
| | | |
|--|--|--|

Actividad Nro. 2 - Recopilar los datos

- Entre todos, elegimos una estrategia para recopilar los datos
 - ¿Armamos un google form para recopilarlos?
 - Hay más opciones, por ejemplo papel y lápiz, y luego pasar a una planilla para procesar.
- Escribir las decisiones tomadas (en cuanto a los casos especiales)

Actividad Nro. 2 - Consigna

¿Cómo vamos a mostrar la respuesta? (visualización)

1. Conformar grupos de 3 integrantes
2. Pedir al docente que comparta los datos recolectados
3. **Mostrar de manera visual la respuesta al objetivo.**

Objetivo: Conocer cuál fue el transporte más usado (por los alumnos de nuestro curso) el día de hoy para llegar a Ciudad Universitaria. En caso de haber utilizado más de un transporte responder con respecto al último utilizado.

| | | |
|--|--|--|
| | | |
|--|--|--|

Actividad Nro. 2 - Discusión

Veamos algunos resultados

| | | |
|--|--|--|
| | | |
|--|--|--|

Actividad Nro. 2 - Discusión

- ¿Piensan que los resultados obtenidos se pueden generalizar a todos los alumnos de la facultad?
- ¿Es similar si trabajamos con alumnos de las últimas materias?
¿Alumnos de doctorado? ¿Alumnos de Maestría?

| | | |
|--|--|--|
| | | |
|--|--|--|

Actividad Nro. 2 - Discusión

- Con los datos recopilados ... ¿es posible responder alguna de las siguientes preguntas?
 - ¿Cuál fue el transporte más usado hoy (por el curso) para llegar a Ciudad Universitaria (último tramo)?
 - ¿Cuál es el ranking de transportes más usados hoy (por el curso) para llegar a Ciudad Universitaria (último tramo) y cuál es su porcentaje de uso?
 - ¿Es verdad que los más jóvenes casi no utilizan el auto como transporte para llegar a Ciudad Universitaria?

Importante: Explicitar siempre (en los resultados) lo que asumimos al recopilar los datos

| | | |
|--|--|--|
| | | |
|--|--|--|

Actividad Nro. 2 - Discusión

- ¿Qué datos deberíamos recopilar para poder responder a todas estas preguntas?
- Pensar en un CENSO, ¿se pueden pedir nuevamente los datos? ¿Cuál es el costo?

| | | |
|--|--|--|
| | | |
|--|--|--|

Cierre

1. Los proyectos de Ciencias de Datos se pueden dividir en etapas
2. Existen distintas metodologías para hacerlo
3. Pensar bien el objetivo antes de recopilar datos

Tarea

Leer “Funes el Memorioso”, Jorge Luis Borges como para poder contárselo al resto de la clase (se encuentra subido en el campus)

| | | |
|--|--|--|
| | | |
|--|--|--|