**Trabajo Final EEA**

**Pautas para la Realización de los Videos Tutoriales**

Objetivo:

El objetivo de este Trabajo Final es la **realización de una serie de videos tutoriales en los que se explique y aplique alguna metodología de la Ciencia de Datos, teniendo en cuenta el enfoque estadístico que se le da a la materia**. Se podrá elegir cualquier tema de análisis cuantitativo aplicado a un conjunto de datos y deberá quedar clara la hipótesis de investigación u objeto de análisis desde el inicio de los videos explicativos.

Los datos podrán ser reales o simulados. El material deberá contener una explicación (al menos intuitiva) del funcionamiento de la metodología. Asimismo, deberá incluirse un caso de aplicación a un conjunto de datos mediante el lenguaje R. El código (script/notebooks) deberá alojarse en Github, para ponerse a disposición de cualquier persona interesada.

Los videos serán publicados en un canal público de youtube, salvo que se solicite lo contrario por la/el estudiante. En la descripción de los videos se incorporará el link de acceso a Github, así como una dirección de correo electrónico para que cualquier persona pueda contactarse.

Modalidad:

* **Grupos**: el trabajo deberá realizarse en grupos de **2 (dos) personas**, de modo tal que ambos integrantes participen activamente en los videos tutoriales.
* **Entrega:** Los videos serán entregados compartiendo una carpeta en google drive por correo electrónico. El código (scripts/notebooks) deberá alojarse en Github, para ponerse a disposición de cualquier persona interesada.
* **Fecha de entrega**: 4 de diciembre de 2021.
* **Temática seleccionada**:
  + En caso de que sea un caso de aplicación, se requiere explicar problema o hipótesis de investigación y metodología de trabajo.
  + En caso de que sea un desarrollo teórico o explicar alguna técnica, sugerimos emplear una demostración en código al menos con datos simulados para lo que se busca demostrar/explicar.
* **Videos**: se sugiere iniciar el video con una pequeña introducción de lo que se verá, explicando el problema que se buscará resolver o la metodología que se busca explicar, luego desarrollar el tema y, por último, hacer un pequeño resumen o mencionar una conclusión del tema o problema de investigación. Lo que no puede faltar en los videos:
  + El título de lo que se va a explicar
  + El saludo o la presentación de la persona que habla
  + El objetivo que va a conseguir la persona que el video tutorial
  + El desarrollo del contenido del tutorial
  + El resultado o conclusión/resumen de lo que se ha enseñado
* **Recursos adicionales**: El trabajo podrá contener recursos adicionales, como ser textos o presentaciones que se depositarán en el mismo Github en el que se hallan los videos.
* **Fuentes externas**: Para la realización del material podrán utilizarse recursos de terceros (gráficos, imágenes, textos o videos) con sus debidas citas/referencias.

Evaluación

Este trabajo representa un tercio de la nota final y será evaluado en una escala de 0 a 100 puntos.

Criterios de evaluación

Serán evaluados el desarrollo e interpretaciones, como también la claridad en la exposición, denotando un entendimiento sobre la temática abordada y la ejecución del análisis. Los objetivos generales que se evaluarán son:

* Claridad en la explicación de la hipótesis y la metodología a desarrollar
* Fundamentar y explicar los procedimientos y decisiones
* Demostrar conocimiento de las técnicas y modelos requeridos
* Desarrollar y comentar su código (en los scripts)
* Entregar material de forma prolija para que pueda ser interpretado por cualquier persona no necesariamente entendida en la temática.

Lista de pautas mínimas a cumplir:

* **Hipótesis de Trabajo**: Al inicio de los videos debe dejarse bien claro cuál es el objetivo del análisis.
* Entrega de videos por google drive.
* Entrega de link de Github incluyendo el código y material complementario si hubiere
* Entrega de dirección de correo electrónico para poder ser contactados en caso de estar de acuerdo con compartir el contacto.
* Duración Mínima de los videos: 25 minutos.
* Duración Máxima de los videos: 50 minutos.
* Lenguaje de aplicación: R.

Sugerencias:

* Contemplar la posibilidad de elegir tema y datos que puedan total o parcialmente conformar el trabajo de especialización y/o la tesis de maestría.
* Generar varios videos cortos, en detrimento de un único video largo.
* Utilizar Rstudio.
* Crear un Notebook para la aplicación de la metodología.
* Se estimula el uso de datos reales en caso de que sirvan indistintamente para ejemplificar la metodología.
* Hacer una investigación previa para ver si hubo otros avances o tutoriales ya realizados sobre la temática.
* Crear un guión o punteo de ideas principales a transmitir en los videos.
* Practicar antes de grabar el video.
* Verificar que se puedan escuchar correctamente las voces.

Lista de Temas Globales Potenciales:

Se presenta un listado no exhaustivo de temas para poder desarrollar los videos. Se espera que el contenido desarrollado sea una extensión o profundización de los temas abordados en el curso o bien algún tema no desarrollado pero que presente relación con el enfoque temático de la materia

* Tests estadísticos sobre el modelo de regresión clásico: tests de heterocedasticidad, no linealidad, normalidad
* Modelos de Regresión Robustos
* Modelos Lineales Mixtos (con efectos fijos y aleatorios)
* Modelos de Variables Instrumentales
* Modelos con variables proxy
* Modelos de Diferencias en Diferencias
* GLM: Poisson, Beta
* Regresión por Componentes principales
* Regresión para series de tiempo
* Regresión por cuantiles
* Regresión bayesiana
* Regresión KNN
* Simulación de datos
* CART: Classification And Regression Trees
* Métodos de Smoothing
* GAM: Generalized Additive Models
* PPR: Projection Pursuit Regression
* Interpretabilidad de modelos: Local interpretable model-agnostic explanations (LIME)
* Interpretabilidad de modelos: SHAP (Shapley Additive Explanation)
* Regresión por Multivariate Adaptive Regression Splines (MARS)
* Least Angle Regression
* Partial Least Squares
* El Algoritmo EM (Expectation Maximization)
* Análisis Discriminante Lineal y Cuadrático (LDA y QDA)
* Clasificación por Naive Bayes
* Clasificación por Discriminant Adaptive Nearest Neighbor (DANN)

Lista de Sitios Repositorios de Datos Públicos:

<https://www.datos.gob.ar/>

<https://data.buenosaires.gob.ar/>

<https://www.kaggle.com/>

<https://datasetsearch.research.google.com/>

# Plagio

Entendemos como plagio cualquier copia o reproducción total o parcial de recursos (trabajos, notebooks, videos, posts, etc.) de otra persona o bibliografía externa sin mención de la fuente original.

El plagio de explicaciones e interpretaciones implicará la calificación de 0 puntos del trabajo práctico y, potencialmente, la anulación de otras instancias de evaluación de la materia. Toda otra fuente empleada para el trabajo, deberá ser explícitamente citada.