*RÚBRICA DE EVALUACIÓN DE ENTREGAS DEL PROYECTO FINAL*

**Data Science**

Una **rúbrica** es un conjunto de criterios y normas que sirven para evaluar el nivel de desempeño en una tarea. Es una herramienta de calificación que simplifica la tarea de corrección y permite lograr una evaluación más objetiva y transparente tanto para el evaluador como el evaluado.

Adicionalmente, la rúbrica ayuda a igualar y alinear los criterios de evaluación entre los diferentes docentes y tutores.

**En resumen, la rúbrica:**

1. Define los criterios de evaluación de cada entrega de proyecto final.
2. Describe y detalla los puntos clave que debe tener cada entrega.
3. Facilita y transparenta la corrección de cada trabajo.

*¿CÓMO SE UTILIZA LA RÚBRICA?* 🎦[Ver Tutorial](https://www.loom.com/share/ad487462a0bc424cb10355989ae7abe2)

*PRIMERA ENTREGA DEL PROYECTO FINAL*

**Se debe entregar**:

1. Presentación de la empresa, organización o problema específico.
2. Preguntas y objetivos de la investigación.
3. Conformación del equipo de trabajo.
4. Indicación de la fuente del dataset y los criterios de selección (Data Acquisition).
5. Generación del primer Data Wrangling y EDA, apuntado a sus datos (insights) univariado, bivariado y multivariado.
6. Análisis de componentes principales.
7. Contar la historia de sus datos
8. Filtros aplicados a los datos. Distribución. Dataset final para analizar.
9. Plantear objetivo/s para esos datos.

**Consigna:** En equipos de entre 3 y 5 estudiantes, gestionar un proyecto con las características necesarias para su implementación en el nivel técnico en una organización. Diseñar un modelo analítico de Data Science en Jupyter Notebooks, acorde a la temática a tratar y su respectiva documentación. Tendrán por lo menos una iteración para la optimización del modelo construido. Todas las métricas deberán estar declaradas correctamente.

| **Modelo por bloques de código (5)** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Criterios** | **Nulo**  AUSENTE  **0 PUNTOS** | **Bajo**  CONFUSO  **0,5 PUNTOS** | **En proceso**  FALTA DE PROFUNDIDAD  **1 PUNTOS** | **Correcto**  PROBLEMAS PUNTUALES  **2 PUNTO** | **Óptimo**  CLARO  **2,5 PUNTOS** | **NOTA** |
| Estructura, Presentación y Objetivos de la investigación | -No presenta una estructura acorde y los momentos no se adecúan a lo estipulado.  -No presenta a la empresa, organización o problema específico.  -Los objetivos no permiten identificar el alcance del proyecto  -No realiza las preguntas, objetivos, o hipótesis de la investigación  -No indica la fuente de información  -No identifica los criterios de selección del dataset | -Presenta con errores a la empresa, organización o problema específico, o bien tiene dificultades para encuadrar su alcance  -Algunos objetivos planteados se alejan de poder definir los alcances del proyecto.  -Realiza preguntas o planteos de hipótesis que no guardan relación con los posibles alcances de la investigación  -Apenas justifica la elección de la fuente de datos con alguna identificación vaga, o esta no responde a las preguntas o hipótesis planteadas | -Presenta a la empresa, organización o problema específico pero no define con precisión o completitud su alcance.  -Los objetivos planteados permiten a priori identificar algunos de los posibles alcances del proyecto pero no todos los que serían evidentes.  -Realiza algunas preguntas o planteos de hipótesis que permitan entender los alcances posibles de la investigación  -Justifica superficialmente la fuente de información elegida, solo resaltando algunos de los criterios de selección del dataset | -Presenta adecuadamente a la empresa, organización o problema específico, pero no logra identificar todas las características  -Los objetivos planteados son acordes y permiten identificar la mayoría de los alcances del proyecto, pero no todos.  -Realiza varias preguntas o planteos de hipótesis para entender en parte los objetivos que se plantean  -Justifica la elección de la fuente de información obtenida, pero omite algunos criterios de selección del dataset que no permiten responder a todas las preguntas o hipótesis planteadas | -Presenta una estructura de tres momentos claramente definidos.  -Presenta con exhaustividad a la empresa, organización o problema específico  -Los objetivos planteados permiten con totalidad identificar todos los alcances posibles del proyecto  -Realiza suficientes preguntas o planteos de hipótesis que permiten con facilidad entender todos los alcances posibles de la investigación  -Justifica adecuadamente la fuente de información obtenida, resaltando todos los criterios de selección del dataset para responder a los objetivos planteados, preguntas generadas e hipótesis planteadas |  |
| Aspectos metodológicos | - El notebook posee varias celdas sin correcta ejecución  - El procedimiento del Data wrangling y EDA es desordenado e inapropiado (genera errores en la interpretación)  - No se aplican filtros al dataset (siendo necesarios) | - El notebook posee algunas celdas sin correcta ejecución  -El procedimiento del Data Wrangling y EDA es desordenado y genera errores en los posibles insights del negocio  -Se aplican malos filtros al dataset | - El notebook se ejecuta con errores no fatales  - El procedimiento del Data wrangling y EDA es escueto y no permite hallar todos los posibles insights del negocio  - No se aplican suficientes filtros, o se realizan de un modo que atenta contra el dataset final | - El notebook se ejecuta totalmente sin errores  -El procedimiento del Data wrangling y EDA es ordenado y acorde para la generación de algunos insights presentados  -Se aplican varios filtros adecuadamente pero no todos los necesarios para conseguir el dataset final | - El notebook se ejecuta totalmente sin errores  - El procedimiento del Data wrangling y EDA es sumamente ordenado y acorde para la generación de insights del negocio  - Se aplican suficientes y correctos filtros al dataset |  |
| Storytelling | -No presenta un plan de investigación o este no cumple con las pautas previstas.  -No se expresa la información en secuencia global-intermedio-detalle  -Los objetos, gráficos e imágenes no expresan con la temática, variable o análisis que se quiere presentar  -La presentación de la información es desordenada  -No detalla información con tītulos, subtítulos | -Presenta un plan de investigación que falla con varias de las pautas previstas  -El plan de investigación es demasiado reducido o está enfocado a uno de los objetivos planteados  -No presenta adecuadamente la información en secuencia global-intermedio-detalle.  -Los objetos, gráficos e imágenes no son claros, o no se adecuan entre sí y a la temática planteada.  -No detalla información con tītulos, subtítulos (markdown) correctamente | - No prefigura un plan de investigación lo suficientemente claro y ordenado para poder responder a la/s pregunta/s de investigación.  - El plan de investigación es acotado a algunos puntos u objetivos planteados  - Falla en varios momentos al presentar la información en secuencia global-intermedio-detalle.  -Los objetos, gráficos e imágenes son claros pero no se adecuan entre sí y a la temática planteada  -No detalla toda la información con tītulos, subtítulos (markdown) correctamente o totalmente. | - Puede prefigurar un plan de investigación claro y ordenado para poder responder a algunas de las preguntas de investigación.  -Presenta un buen plan de investigación pero no es lo suficientemente detallado para responder a todas las preguntas planteadas  -La información se presenta preponderantemente en secuencia global-intermedio-detalle  -La mayoría de los objetos, gráficos e imágenes expresan correctamente la temática, variable o análisis que se quiere presenta  -Detalla suficiente información con títulos y subtítulos (markdown) pero no en su totalidad. Ej: no respeta toda la estructura | - Puede prefigurar un plan de investigación claro y ordenado para poder responder a toda/s la/s pregunta/s de investigación.  -Presenta un excelente plan de investigación, el mismo es detallado y se encamina a responder las preguntas planteadas  -La información se presenta en perfecta secuencia global-intermedio-detalle  -Los objetos, gráficos e imágenes expresan correctamente la temática, variable o análisis que se quiere presentar  -Detalla toda la información con tītulos, subtítulos (markdown) correctamente |  |
| Análisis de componentes | -No encuentra, ni detalla insights.  -No reduce ninguna variable del dataset  -No identifica ‘nuevas’ variables que representen a las anteriores  -No presenta el ‘dataset definitivo’ para el modelo a utilizar | -Encuentra insights poco robustos producto de la investigación, o estos no son significativos  -Apenas reduce las variables del dataset, encontrando muchas irrelevantes  -Reduce variables fundamentales en el dataset para el objetivo de la investigación.  -Presenta ‘nuevas’ variables que no representan a las anteriores  -No correlaciona variables entre sí, o lo hace de un modo equívoco  -El dataset presentado no puede ser utilizado como *input* de modelos de *machine learning* sin generar errores graves | -Detalla pocos insights robustos producto de la investigación. Algunos no son significativos  -Reduce una porción de las variables del dataset pero aún se encuentran en la fuente ciertas variables a reducir  -Reduce variables del dataset que son o serían trascendentales para el objetivo de la investigación.  -No identifica suficientes ‘nuevas’ variables  -Realiza algunos análisis de algunas variables  -No logra correlacionar correctamente las variables del dataset  -El dataset presentado aún no puede ser utilizado como *input* de modelos de *machine learning* | -Detalla insights robustos producto de la investigación  -Logra reducir variables importantes del dataset pero omite alguna insignificante  -Reduce o elimina alguna variable del dataset que puede ser importante para el objetivo de la investigación  -Identifica ‘nuevas’ variables pero las presenta de forma confusa.  -Realiza varios analisis de variables individuales pero no logra profundizar en todas  -Correlaciona variables del dataset entre sí omitiendo algunas  -El dataset presentado puede ser utilizado para ser *input* de modelos de *machine learning* pero aún puede generar sesgos evitables | -Detalla insights robustos y fundamentales producto de la investigación.  -Reduce adecuadamente las variables del dataset  -No reduce o elimina variables del dataset que son o pueden ser importantes para el objetivo de la investigación  -Identifica adecuadamente todas las ‘nuevas’ variables que representan a las del dataset original  -Realiza suficientes análisis de variables individuales, y correlaciones entre dos y/o más variables  -Presenta un dataset listo para ser *input* de modelos de *machine learning* |  |

* Aclaración: Los estudiantes pueden elegir entre diferentes temáticas sobre las que generar sus proyectos. Las elecciones pueden ser amplísimas y variadas, lo que se evalúa en esta consigna es la gestión del proyecto.