**Temas faltantes para replicar en este repo:**

- hacer todos los códigos difíciles

Links con cosas faltantes para replicar:

- tablas de contingencia con datos categorizados:

- análisis series de tiempo: <https://github.com/youssefHosni/Practical-Time-Series-In-Python>

- códigos de eda del resto del equipo, visto hasta ahora que recuerde:

- raul bmo: <https://gitlab.com/empresas-cmpc/cd4cmpc/proyectos-analytics/cmpc-data-science/-/blob/master/src/use_cases/balance_mill/increment-08.3.8/data_drift.ipynb?ref_type=heads>

- tiare wetlap: <https://gitlab.com/empresas-cmpc/cd4cmpc/proyectos-analytics/cmpc-data-science/-/blob/master/src/use_cases/Proyecto%20Wetlap/data_discovery/2.model.ipynb?ref_type=heads>

- <https://gitlab.com/empresas-cmpc/cd4cmpc/proyectos-analytics/cmpc-data-science/-/blob/master/src/use_cases/Proyecto%20Wetlap/data_discovery/1.data-profiling-eda2.ipynb?ref_type=heads>

- <https://gitlab.com/empresas-cmpc/cd4cmpc/proyectos-analytics/cmpc-data-science/-/blob/master/src/use_cases/Proyecto%20Wetlap/data_discovery/0.data-profiling-eda.ipynb?ref_type=heads>

**Ideas BMO:**

- revisar datos faltantes (nulos)

- revisar datos faltantes por timestamp cada 5 minutos. Pd.range(start, end, step)

- revisar frecuencia de actualización de la medición para ver datos de laboratorio. Value(t) – value(t-1) es igual = 0, varía = 1

- ydata-profiling (pandas profiling) para ver perfilamiento básico y alertas automáticas que despliega sobre los datos

- análisis clásico de dispersión, histogramas, correlaciones

- categorizar variables por algun criterio, ej producción niormal y producción planta reducida

- comparar histogramas de las variables, correlaciones, que tanto varian los datos bajo esas condiciones