

Taller 1

ivantrujillo1229@gmail.com

1 Punto 1

- 1) En el siguiente link encontrará la base de datos de instacart (**click aquí**)
deben usar las bases de datos:

- order_products_train
- products

Definan una heurística para seleccionar las primeras 5 reglas, que puedan aportar al negocio de **instacart**.

puede ser de ayuda los métodos **merge** y **replace** de las clase **DataFrame** de la librería pandas.

2 Punto 2

En el siguiente link encontrará el libro **The elements of statistical learning** (**Click aquí**) realice la simulación del apartado 7.10.2 *The wrong and Right Way to do Cross-validation* (página 245).

Para este punto puede ser útil revisar la documentación de:

sklearn.model_selection.StratifiedKFold y **scipy.stats.pointbiseri**