# \*\*README\*\*

---

## Table of contents

1. General info

2. Phase 1

3. Phase 2

4. Phase 3

---

### \*\*General info\*\*

> Project 4 consists of various types of data analysis for an online book shop called "Rester Livres"

> We will be analyzing our data using: central tendencies, univariate, bivariate, and finding a correlation between multiple variables using Pearson's coefficient correlation, Chi, p\_value, and contingency.

> I also provide data visualization with the help of different graphs and tables to provide easier data reading.

### \*\*Phase 1\*\*

> Our first step is to import different Python libraries that will enable us to call different functions to execute our codes

> Then we will start importing our initial data sets, which are :

- **\*\*\*Products\*\***\*

- **\*\*\*Customers\*\***\*

- **\*\*\*Transactions\*\***\*

### \*\*Phase 2\*\*

> Second step brings us to start investigating our initial data sets and cleaning

> We will encounter some data to drop, add columns, impute and fix data types before we can merge our 3 data sets into our one main data frame that we call **\*\*\*TPC\*\***\*

### \*\*Phase 3\*\*

> Data analysis starts with central tendency and univariate analysis of our initial data sets individually

> Then we will go further by dong cross-analysis between different variables

> We will finalize our analysis by finding relationships between variables, so we will be able to check if there is a correlation between them and how they affect each other

> We will end up with a keynote presentation to summarize the project and also suggest some recommendations to improve algorithm; hence sales and revenues