

# Welcome to the Product Analytics Technical Test: delivery app

## What we value:

- • A clean coding style
- • Efficient solutions to the problems given
- • Idiomatic use of Python
- • Appropriate use of Python's built-in functions and standard libraries

## 1 Exploratory Data Analysis (Statistical Programming Exercise)

You've been given a **data set** about orders coming from Fake partners in the app. Fake partners are the stores that are not integrated with our Delivery app directly so our content team manages their product catalog and prices for them. Fake partner orders are charged to the customer upon delivery and in many cases there is a mismatch between the total amount at checkout in the app (*products\_total*) and what the courier pays at the store (*purchase\_total\_price*) causing many problems. When the *products\_total* is lower than *purchase\_total\_price* we call them **under-authorized orders**, otherwise is a correctly authorized order. We want to move away from *charge-on-delivery* to an *authorize-and-capture* model but we first need to understand the price fluctuation of past orders to know the risk of doing so.

Dataset description:

- - order\_id
- - activation\_time\_local: local time when the order was activated
- - country\_code
- - store\_address
- - final\_status
- - payment\_status
- - products: number of products in the order
- - products\_total: total amount at checkout (€)
- - purchase\_total\_price: amount the courier paid at the store (€)

Your task is to perform an EDA process (Python) with this data to answer the following questions:

1. What percent of orders are under-authorized?
2. What percent of orders would be correctly authorized w/ incremental authorisation (+20%) on the amount at checkout?
3. Are there differences when split by country?
4. For the remainder of orders that would be outside of incremental auth what values would be necessary to capture the remaining amount?
5. Which stores are the most problematic in terms of orders and monetary value?
6. For under-auth orders is there a correlation between the difference in the prices and

the cancellation of the order? In other words: Is an order more likely to be cancelled as the price difference increases?

### **Team up with your classmates**

1. Create a private repository on github
2. Add each other as collaborators within the repository
  1. If you're working collaboratively, watch out for git conflicts
3. Create a **main.py** file that when ran, it executes the functions that solves each of the questions: include visualizations.
4. Create a **presentation** with your findings and present to your class
5. Prioritize responding questions but do take care of details if possible: README.md file, documentation of code, etc.

There is a time constraint, so respond to as many questions as you can.