# Product Analyst – Personalization & Relevance

Case Study

## Problem Definition

We have launched our product in the South African market, and our core goal as a team is to maximize buyers’ chances of finding the content they are looking for and being able to buy it.

Today we will be focusing on the latter, namely, **how can we maximize buyers’ chances of closing a deal after making a reply on an ad?**

## Context

**OLX Group** provides Online Classifiedsplatforms, through which sellers (a.k.a. Listers) post goods and services for other users (buyers) to find. When a user finds an id he or she is interested in, they can contact the seller through multiple channels we collectively refer to as reply channels, and each act of contacting a seller is defined as a **reply.**

For the sake of this analysis, we are focusing exclusively on our largest channel: **chat replies.**

A chat reply is a unique conversation between a buyer and a seller regarding a specific item. They can be classified as successful (when the seller replies back) and unsuccessful (when the seller doesn’t reply back) and we can use their length (total count of messages) to estimate the relevance of the conversation.

## Task

**Recommend a product strategy to maximize buyers’ chances of closing a deal after sending a reply.**

You may take any approach you like, choose the target metric you prefer, and use any tool you’re comfortable with (excel, Python, R). The dataset is small enough to be handled by excel.

Your end deliverable should be **a product recommendation** that, in your opinion, makes sense in our context and is reasonably doable.

Please send back your recommendation, rationale, and any data analysis you perform to get to it.

# Appendix: Dataset Definition

You are given two datasets:

* item\_data.csv – Data on each of the items listed in our platform that received a reply in the last week
* category\_data.csv – Reference of category names and id’s

### Item Data

Field description for the dataset in item\_data.csv

|  |  |  |
| --- | --- | --- |
| Field name | Data Type | Description |
| date\_live\_nk | date | Date when the ad was posted and live on the platform (i.e. findable in search results) |
| category\_sk | string | Surrogate key (i.e. unique key) identifying the category the item is posted in |
| listing\_sk | string | Surrogate key (i.e. unique key) of the item itself |
| user\_sk\_encrypted | string | Surrogate key (i.e. unique key) of the seller (encrypted through md5 hash) |
| replies | integer | Amount of unique conversations started by buyers, with respect to the item, during the week preceding September 21st |
| successful\_replies | integer | Amount of unique conversations started by buyers during the week preceding September 21st to which the seller responded back |
| average\_conversation\_length | integer | Average length, in number of messages, of all the conversations that buyers generated on the item during the week preceding September 21st |
| first\_reply\_date | date | Date when the first reply was generated for the item |
| last\_reply\_date | date | Date when the last reply was generated for the item |

### Category Data

Field description for the dataset in category\_data.csv

|  |  |  |
| --- | --- | --- |
| Field name | Data Type | Description |
| category\_sk | string | Surrogate key (i.e. unique key) identifying the category |
| category\_l1\_name\_en | string | English name of the parent catgory (e.g. Vehicles) |
| category\_l2\_name\_en | string | English name of the category (e.g. Cars) |