**Shopify App Analysis**

You have been hired to review the landscape of apps on the Shopify platform, using data scraped from publicly available Shopify websites. You want to figure out what key factors play into the success of a Shopify app.

For each numbered part in this project, you should produce a page in a Power BI report. Each question itself should be a visualization. For each subquestion, take a screenshot of your entire screen to show your work.

**The Data**

The shopify.xlsx dataset contains public data scraped from the [Shopify App Store](https://apps.shopify.com/). It includes 4 tables:

* apps: Details of the apps on Shopify apps marketplace
* apps\_categories: Join tables to connect apps with categories
* categories: Categories of the apps. Each app has multiple categories
* reviews: Each review (row) contains information on user opinion about the related app (rating and comment). Also, it contains the response from the developer if present.

**Part 1: App Landscape**

First, we’ll find key statistics on the types of apps that are out there. This question uses the Apps table. Make a new sheet in your Power BI (.pibx ) file called App Landscape for this question. Remember to take a screenshot for each question answered.

1. Make a KPI Card (a visual with a single number) that counts the unique number of apps. Add it to your App Landscape sheet.
2. Make a Line Chart getting the sum of the review count on the Y-axis, and the lastmod date on the X-Axis. Hint: lastmod date should not be in “date hierarchy” format for your X-Axis.
3. Make a Scatterplot with the reviews\_count on the X axis and the average rating on the Y axis. Annotate your interpretation of this with an inserted Text Box next to this scatterplot.

**Part 2: Reviews**

Make a new sheet in your .pibx file for this new section, named Reviews. Remember to take a screenshot for each question answered.

1. Create a new Column in the Reviews table named helpful\_reviews using a DAX expression, which multiplies the rating by 1+helpful\_count to weigh the reviews by how helpful they’ve been found. The formula in mathematical form would be rating \* (1+helpful\_count). Make a Card with the average value of the new helpful\_reviews column.
2. Create a new Column in the Reviews table named developer\_answered using a DAX expression, which is 1 (or TRUE) if the developer\_reply column is not blank and 0 (or FALSE) if the column row is blank. Make a scatterplot comparing the average rating on the Y-Axis by the value of the developer\_answered column on the X-Axis.

**Part 3: App Reviews**

Make a new sheet in your .pibx file for this new section, named App Reviews. Remember to take a screenshot for each question answered.

1. In the data model, create a new relationship between the Reviews table and the Apps table. Use the app\_id column from the Reviews table, the id column from the Apps table. Make the relationship many-to-one built as many (Reviews table) to one (Apps table). Using this new table, make a bar chart with the developer on the X-Axis, and the sum of rating on the Y-Axis.
2. The chart in the previous question is misleading because an app could have a high sum of rating because it has many one-star reviews. Make a new bar chart with developer on the X-Axis against the helpful\_review average on the Y-Axis.
3. Which developers are the most responsive? Make a bar chart with the developer from the apps table and the developer\_answered column you created in question 2.2. Add a Filter for this visual only which selects only the rows where reviews\_count is greater than 500.