Week 4 | Examination assignments

Data visualization

Erik Vanhatalo

2025-03-16

Contents

Regarding the Assignments

This document outlines the **individual** examination assignments for this week. We strongly recommend that you work through the training assignments for this week before progressing to the examination assignments.

Prepare and document solution before the examination session

You need to prepare your solutions before the examination session. At the examination session, you will need to have your imported datafiles and prepared solutions easily available so that you will be able to perform the tasks you are given at the session.

Quarto - recommended way of documenting your solutions

There are some options for you to prepare, document and save your solutions. You can do it in a standard R script (.R file). which is easy and straightforward. However, we strongly recommend that you use Quarto because it makes it possible for you to combine executable code with your own notes. You can also create *MS Word*, *pdf* or *html* documents with your solutions that you can use as you prepare. During Week 1 there will be an introductory lecture featuring Quarto. Quarto is integrated in RStudio. **Whatever choice you make it is important that you are able to run your prepared solutions fast and easy at the examination sessions.**

While Quarto is very good to document your R code and visualizations you produce, for Week 4 you also need to **find a way to prepare and document your Power BI Dashboard** that you develop.

1. Examination assignment 1 - Power BI Dashboard for "Global superstore"

In this assignment you will work with **"global_superstore_2016.xlsx"** dataset that you can find on Canvas. This dataset is based on a simulation of retail sales. Data come from stores in multiple countries. The dataset includes customer information, orders, products, dates, prices, profits etc. The simulation data have many observations and many features and is therefore excellent for practicing exploration of retail sales data. **The most interesting variables are in the orders sheet including a total of 51 290 orders.** The table below lists the variables/features and brief explanations.

Table 1: List of variables/features with brief explanations.

Variable/Feature	**Explanation**
Order ID:	A unique identifier for each order.
Order Date:	The date and time the order was placed.
Ship Date:	The date and time the order was shipped.
Ship Mode:	The method used to ship the order (e.g. standard, express).
Customer ID:	A unique identifier for each customer.
Customer Name:	The full name of the customer.
Segment:	The customer segment such as Home Office or Corporate.
Postal Code:	Postal code of the customer's residence, if available.
City:	The city where the customer resides.
State:	The state where the customer resides.
Country:	The country where the customer resides.
Postal Code:	The postal code of the customer's residence.
Region:	The geographic region where the customer resides.
Market:	Market name or abbreviations.
Product ID:	A unique identifier for each product.
Category:	Broad product category, such as Furniture or Technology.
Sub-Category:	Specific product sub-category, such as Chairs or Phones.
Product Name:	The name of the product.
Sales:	The total sales revenue for the product.
Quantity:	The number of units of the product sold.
Discount:	The discount applied to the product, given such as 0.1 = 10%.
Profit:	The total profit, in $, earned from the product.
Shipping cost:	The cost, in $, of shipping the order.
Order priority:	A classification of how prioritized the order was.

1.1 What your Dashboard needs to include

Note: You should approach the tasks below with the aim of creating a Dashboard that can be a tool to help management stay informed, find improvement areas, and make better tactical and strategic decisions for the "Global Superstore" going forward.

Import the dataset into MS Power BI and create your own Global Superstore Dashboard. Your job here is to **be creative** and develop your own Dashboard that focuses on what you think is most interesting!

In your Dasboard you should at least:

- Change the background of your Dashboard to something nice-looking and add text explanations for what the Dashboard is about.

- Include **at least five different** visuals types of the built-in visuals in Power BI.

1.2 Guidance for designing your Dashboard

You should **be creative** with your Dashboard, but as **a guidance** we provide some examples of broad analysis "components" or question that could be of interest in the sub-headings below. *Please note that the above examples are just a way to get you thinking creatively!*

1.2.1 Sales, costs, profit over time. Examples.

- Be able to follow sales, shipping costs, and/or profit over time.
- Be able see total profit, shipping costs, and/or profit on a yearly basis.

1.2.2 Product, order, and shipping analysis. Examples.

- Visualize which product category and/or sub-category is most profitable?
- Visualize which product category and/or sub-category has most sales?
- Visualize which is the most profitable product?
- Visualize which order priority that is most common?
- Visualize which product category and/or sub-category has been most discounted?
- Visualize which is the most common shipping-mode?

1.2.3 Market and customer analysis. Examples.

- Visualize **on a map** where customers reside and where the most profit is made.
- Visualize which market is most profitable?
- Visualize which country is most profitable?
- Visualize which city has most sales?

2. Examination assignment 2 - Integrate advanced R visuals using 'ggplot' in your Dashboard

In this assignment you are supposed to show that you can integrate advanced R visuals in the Power BI Dashboard. You should work with the 'ggplot2' package (Wickham 2016) within the 'tidyverse' family of packages (Wickham et al. 2019). You should continue to work with the **"global_superstore_2016.xlsx"** dataset.

Note: You may have run out of room on your Dashboard in Assignment 1. You can add pages to your Dahboard where you integrate R visuals.

2.1 Tasks to complete

1. Create **at least two** advanced visualizations using the 'ggplot2 package' and integrate them on your "Global Superstore" Dashboard. You should create visualizations that are not available (not possible) in the standard Power BI visuals.
2. **Explain** what you are trying to show with your R visualizations and document your R code that produces them.

2.2 Guidance and tips

We recommend that you develop your visuals in RStudio first testing things out. This also allows you to effectively document your R code for producing the visual. We advise you to document your R visual using Quarto and then take the code and integrate it into your Dashboard.

3. The new assignments at the examination session?

If you have solved the above tasks, documented your solutions, and are able to run the code fast in the classroom you should be well-prepared to solve the additional assignment(s) that you will get at the examination session. In preparation for the examination assignments it is also important that you and have worked through the training assignments. If you have done this, **do not worry!**

References

Wickham, Hadley. 2016. "Ggplot2: Elegant Graphics for Data Analysis." https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the {Tidyverse}" 4: 1686. https://doi.org/10.21105/joss.01686.