**Introduction to Data Visualization   
& Infographics**

**Thomas More Geel**

**What is Data Visualization and its Purpose?**

*Big data* has been one of the most overused buzzwords in the last couple of years within the IT industry despite the fact that a true, uniform definition for the term has still not been found. It may become difficult to talk about the capturing and managing processes of big data, but even more challenging may be to convert this huge amount of data into something meaningful for the data owner. If the purpose of business intelligence is to enable informed decision making for stakeholders with the use of its own resources (such as data), how can we effectively present our data insights to them? The practice of data visualization helps us to present data in a graphical, yet clear format utilizing our knowledge of how the human brain processes graphs, colors, size and shape.

Most importantly however is to enable the stakeholder to understand complex data gathered from big data processes to understand patterns, relationships or comparisons to seek out the next best strategy move. Some people may call data visualization a form of science due to the heavy analysis required before the data is fit for purpose; others may call it a form of art due to the freedom of choice as of how you would want to format the charts and graphs required. The truth may be somewhere in the middle: having a deep understanding of the data through analysis enables you to think in different ways of how your insights could be presented.

The question now is why data visualization is gaining more attention while it does not sound as an entirely new practice. On the contrary, examples of data visualization methods can be found as early as the seventeenth century. The big game changer in today’s industry is that data visualization is shifting from *static visualization*, capturing data from one point of time into one single format, to *dynamic visualization*. With the aid of more user-friendly interfaces and interaction tools for users such as drill-down and filtering options, data can be presented dynamically in real-time. Think of quickly filtering through years, market segments, product categories, enabling stakeholders to answer multiple business questions in relatively less time. Having knowledge of both data practices and business strategy knowledge allows you to not only understand *what* data needs to be present but also a good idea of *how* it should be presented as well.

**What are infographics?**

A good infographic is worth a thousand words. An infographic, which is the graphical representation of information, allows you to communicate data and conceptual information more effectively than written reports. Because humans process visuals much faster than text, infographics are a great way to deliver potentially complex information while keeping your audience’s interest. Like data visualization the use of visual shapes, colors and different sizes helps our brain to process information in a smarter way.

However, the design process of an infographic requires more than just drawing nice visuals and using bright colors. Of course, it is important to know and apply good design practices but it is even more important to know your audience and their information needs. When you can also add good storytelling to it, that’s when you start to make real impact. Therefore, making infographics requires you to let work both your analytical and creative skills together.

Infographics are often used to visualize data, step-by-step processes, timelines, research findings, organizational hierarchies, and many other types of information.

**What is Tableau Desktop Software?**

Tableau Desktop is software that promotes *live, interactive* visualization of data using a simple drag-and-drop system of dimensions and measures pulled from the used data source(s). Tableau is able to connect to data from a wide variety of sources, whether it would be a single Excel spreadsheet or a data log pulled from Salesforce. Its statement of providing more dynamic data analysis is supported by the software’s ability to provide dashboards with multiple charts and graphs being able to respond and change their presentation based on the user’s actions: whether it is hovering or clicking on a single point in the view or actively using filters available to them, Tableau allows for interaction within the dashboard so stakeholders can easily select the figures they are looking for.

But data visualization should not stop at *what* someone is looking for: one should also dare to ask the question *why* the figures are the way they are. Besides simple data visualization through charts and graphs, Tableau Desktop is also equipped with statistical functions to better explain trend lines, regressions and correlations between variables. Because data visualization within Tableau Desktop can be achieved at a relatively fast pace due to its drag-and-drop controls, analysts are now given more time and possibility to explore the data further to further support the business in enabling informed decision-making. Tableau Desktop also allows for *story-telling analysis* by letting users create snapshots of their insights and highlighting those they wish to explain further to stakeholders if needed, allowing for narratives combined with interactive data graphs that would otherwise not be possible.

Finally, data visualization and overall presentation in Tableau Desktop does not stop at the level of simply looking at a chart or graph. Users of the software are also able to enhance the *user experience* of the produced worksheets, dashboards or stories by customizing the design of both the worksheets and the available interactive data filters. Furthermore, they are able to include and customize the format of other small details such as the *tooltip*: an annotation field that pops up when you hover over a data point in the view. With these functionalities included, those being assigned to create data visualizations using Tableau Desktop for large groups of stakeholders are now given much freedom to realize their demands in creative and intuitive ways.

**The market of Data Visualization in 2017 by Gartner**



Feb 2017

Feb 2018



**Assignment**

**Content:**

File type: .csv

File size: 303kb

This dat aset contains data about the 1000 most popular movies from the last 10 years (2006-2016) on IMDB (www.imdb.com). The Internet Movie DataBase is an online platform where users can find information about movies and series. Users can also vote and add movies to their online library.

**Business Questions:**

Try to provide a meaningful answer to the following questions by visualizing them in an unambiguous way. Reports/dashboard are created using the Tableau Desktop software.

***Don’t hesitate to display extra insights you find important!***

* Top-10 movies (based on revenue) per:
  + Genre
  + Year
  + Genre & Year
* Top-10 movies (based on rating) per:
  + Genre
  + Year
  + Genre & Year
* Longest and shortest runtime movie per:
  + Genre
  + Year
  + Genre & Year
  + Actor
  + Director
* Most successful movie (based on revenue) per:
  + Actor
  + Director
* Most successful movie (based on rating) per:
  + Actor
  + Director
* Most successful year (based on revenue) per:
  + Actor
  + Director
  + Genre
* What is each director’s favorite:
  + Actor to work with (based on frequency)
  + Genre to produce films in (based on frequency)
  + Combination Actor & Genre
* Make a comparison between the 5 best rated movies from two categories. (Dashboard)
* Actors working in those movies
* Directors working on the movies
* Ratings of the movies
* ‘KPI dashboard’
* Which actors work together the most in badly rated movies? (*hint: LOD*)
* Find out the worst rated movies
* Find out which actors worked in these movies
* Find out which actors work alongside each other in these movies
* Which genre is the preferred one for each actor?
* Find out in which genre each actor has worked the most often
* Which actor is rated best in each genre? (*hint: Table Calc*)
* Find out for each genre which actors worked in the best rated movie, if more than one find out which actor has worked in the second best rated movie, if more than one in both, find out…….
* [Bonus] Storytelling: “*Six Degrees of Kevin Bacon*”
  + Show how Kevin Bacon can be linked to any movie, actor or director in as few steps as possible.

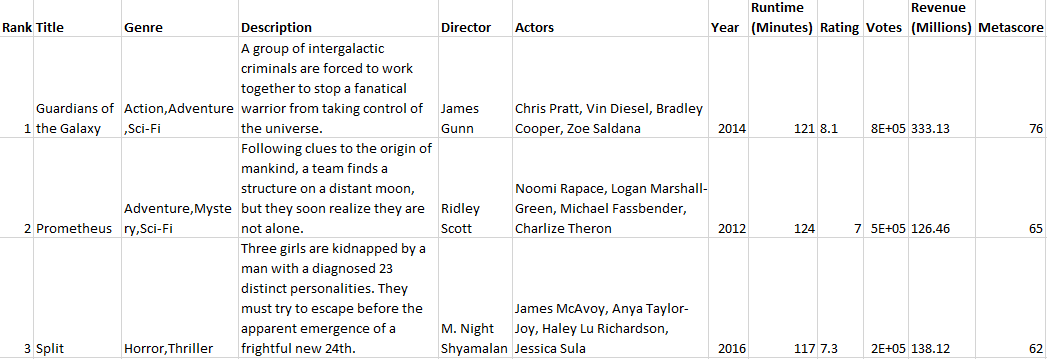
Example: Start with the movie *The Wolf of Wall Street*.

* + - Matthew McConaughey played in *The Wolf of Wall Street*.
    - Christopher Nolan worked together with Matthew McConaughey on a movie (*Insterstellar*).
    - Christian Bale played Batman in Christopher Nolan’s Batman-trilogy.
    - Jennifer Lawrence and Christian Bale both acted in the movie *American Hustle*.
    - **Kevin Bacon** and Jennifer Lawrence acted in the movie *X: First Class*.

**Available Fields:**

* Rank: ranking of overall titles
* Title: movie title
* Genre: genre of the movie. A movie can have multiple genres
* Description: short description of the movie
* Director: movie director
* Actors: actors playing in the movie (4 per movie)
* Year: release year
* Runtime: movie duration in minutes
* Rating: IMDB rating
* Votes: amount of votes
* Revenue: Total revenue in millions
* Metascore: Scores given by movie critics

**File Preview:**



**Awards:**

* Most Creative Thinking
* Best Individual Visualization
* Best Technical Solution
* Best Infographic
* Best Overall