



AI

VISUALIZATION TOOLS

ENHANCED INSIGHTS  
PRESENTATION

# introduction

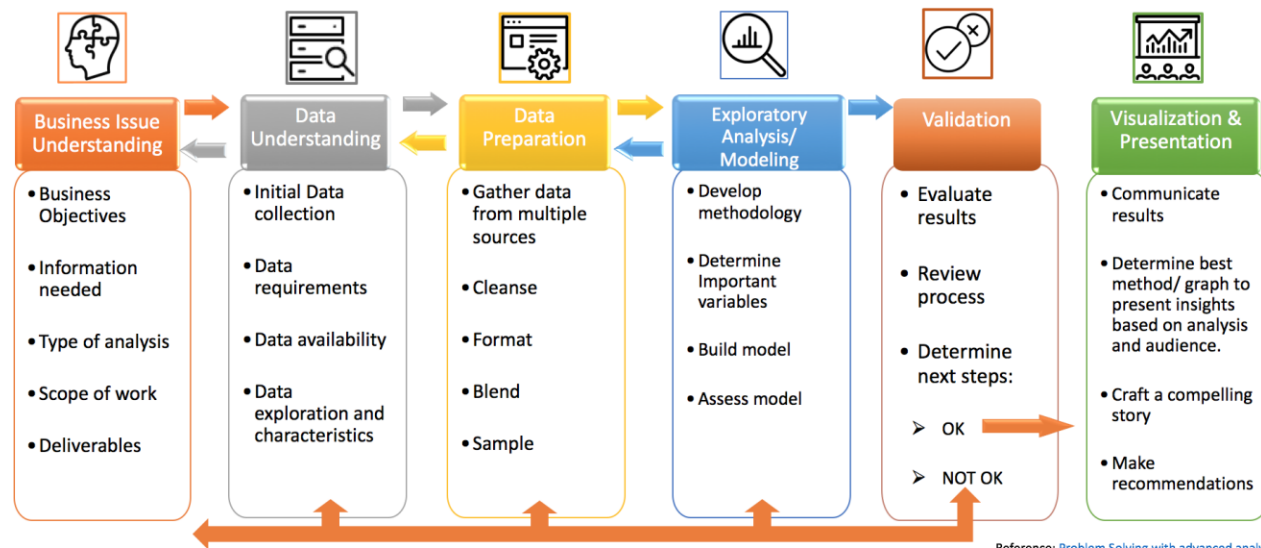
## 6.1 Introduction

While designing a data analytics project, we are often left wondering where to begin with in the first place? From data collection, cleaning, exploration, analysis and visualization, there is a lot that needs to be done in order to derive an insight that is - actionable & profitable, for the business.

There seems to be a no set way to approach this problem. However, in order to provide a framework to organize the work needed by an organization and deliver clear insights from data, it's useful to think of it as a cycle with different stages. ("Big Data Analytics - Data Life Cycle," [n.d.](#)). This article explains a data science framework, breaking it down and taking us through each step of the project lifecycle to get us familiarized with the whole process in a simpler way. ("HOW DO I START A DATA PROJECT: UNDERSTANDING THE LIFECYCLE OF A DATA ANALYSIS PROJECT" [2019](#))

# Life cycle of a data analysis

Life Cycle of a data analysis project based on  
CRISP- DM methodology



Reference: [https://en.wikipedia.org/wiki/Cross-industry\\_standard\\_process\\_for\\_data\\_mining](https://en.wikipedia.org/wiki/Cross-industry_standard_process_for_data_mining)  
Reference: [Problem Solving with advanced analytics](#)

# Steps of visualization

## 6.1.1 Step 1: Understanding the Business Issues

At the start of the project, the focus is to get a clear understanding of the overall scope of the work, business objectives, information the stakeholders are seeking, the type of analysis they want you to use, and the key deliverables. Defining these elements prior to beginning the analysis is important, as it helps in delivering better insights. Also, it is important to get a clarity at the beginning as there may not be another opportunity to ask questions before the completion of the project.



## 6.1.2 Step 2: Understanding Your Data Set

This phase starts with an initial data collection and proceeds with activities like data quality checks, data exploration to discover first insights into the data, or to detect interesting subsets to form hypotheses for hidden information. There are a variety of tools we can use to understand the data. Depending on the size of the dataset, we can use Excel for manageable datasets, or use more rigid tools like R, Python, Alteryx, Tableau Prep or Tableau Desktop to explore and prepare the data for further analysis.



### 6.1.3 Step 3: Data Preparation

Once the data has been organized and all the key variables have been identified, we can begin cleaning the dataset. Here, we will handle missing values (replace with means, drop the rows or replace with the most logical values), create new variables to help categorize the data, and remove duplicates. Data preparation tasks are likely to be performed multiple times, and not in any prescribed order. After this step, the final dataset is ready to be fed into a modeling tool for further analysis.

From a business perspective, throughout the data preparation process the need is to develop an ever-increasing understanding of the data's structure, content, relationships, and derivation rules. It is imperative to verify that the data exists in a usable state, and its flaws can be managed, and understand what it takes to convert it into a useful dataset for reporting and visualization. In such a scenario, leveraging Data profiling can help explore the actual content and relationships in the enterprise' source systems. Data profiling can be as simple as writing some SQL statements or as sophisticated as a special purpose tool. Tableau's Data Prep for instance is a great tool for profiling data for small scale projects. With enterprises, many ETL vendors offer a variety of tools can be chosen based on the need and budget of the business.



### **6.1.4 Step 4: Modeling**

In this step, we will use various modeling techniques to test the data and seek out answers to the given objectives. Typically, there are several techniques for the same data mining problem type, with some specific requirements on the form of data. Common models include linear regressions, decision trees, and random forest modeling, among others.

### **6.1.5 Step 5: Validation**

Once we are done building the model (or models) and proceed to the final deployment, it is crucial to assess the model thoroughly and review the steps executed to construct the model, to ensure that it properly achieves the business objectives. Did the models work properly? Does the data need more cleaning? Did you find the outcome the client was looking to answer? If not, you may need to go over the previous steps again. You should expect a lot of trial and error!



### 6.1.7 Step 7: Documentation

An important addition to the steps in a data visualization project is documentation. Similar to the projects done in class, this documentation should provide a brief description of the project, data sources, data profile and quality, limitations of the data or arising during the use of data, key transformations and models introduced and their impact or usefulness in enhancing the quality of visualization. Lastly, this documentation should also note the issues encountered when working with the data or creating a certain visualization that can be addressed in the future.



### 6.1.6 Step 6: Visualization

Creation of the model is generally not the end of the project. Even if the purpose of the model is to increase knowledge of the data, the derived information will need to be organized and presented in a way that is useful to the customer. Depending on the requirements, this step can be as simple as generating a report or as complex as implementing a repeatable data scoring (e.g. segment allocation) or data mining process.

In many cases, data visualization will be crucial in communicating your findings to the client. Not all clients are data savvy, and interactive visualization tools like Tableau are tremendously useful in illustrating your conclusions to clients. Being able to tell a story with your data is essential. Telling a story will help explain to the client the value of your findings.

As with any other project, it is important to identify the business objectives clearly. Breaking the process into steps will ensure we get the best deliverables for our clients.



# How to structurew od data viz project

## 6.6 How to Structure a Data Viz Project

### 6.6.1 Hook The Audience on Your Key Message

A Hook is simply something that draws the reader in. Hooks communicate to the reader right away why he/she should pay attention to your project. They can be surprising, unusual, or controversial, but one thing's for certain: hooks can't be ordinary. After all, if your project tells people things they already know, then what does your project do for them?

To help you formulate your hook, ask yourself: what is the biggest takeaway? If someone could only remember one thing from your project, what would it be? This takeaway could be a novel point of view, startling new research, or a bold opinion. Often times, this is going to be the main argument of your project. Condense your argument into a short, digestible, memorable chunk.

Journalists have to write hooks all the time: their hooks are called headlines. Nowadays headlines have a bad reputation, but they've been a part of journalism since the beginning for a reason: they work. A great headline does all the things a great hook does: it summarizes the key information of the article, and it attracts the reader's attention so that he/she will read the article. This isn't to say that you should publish click-bait, but it's important to entice your reader's attention no matter what kind of project you're doing.

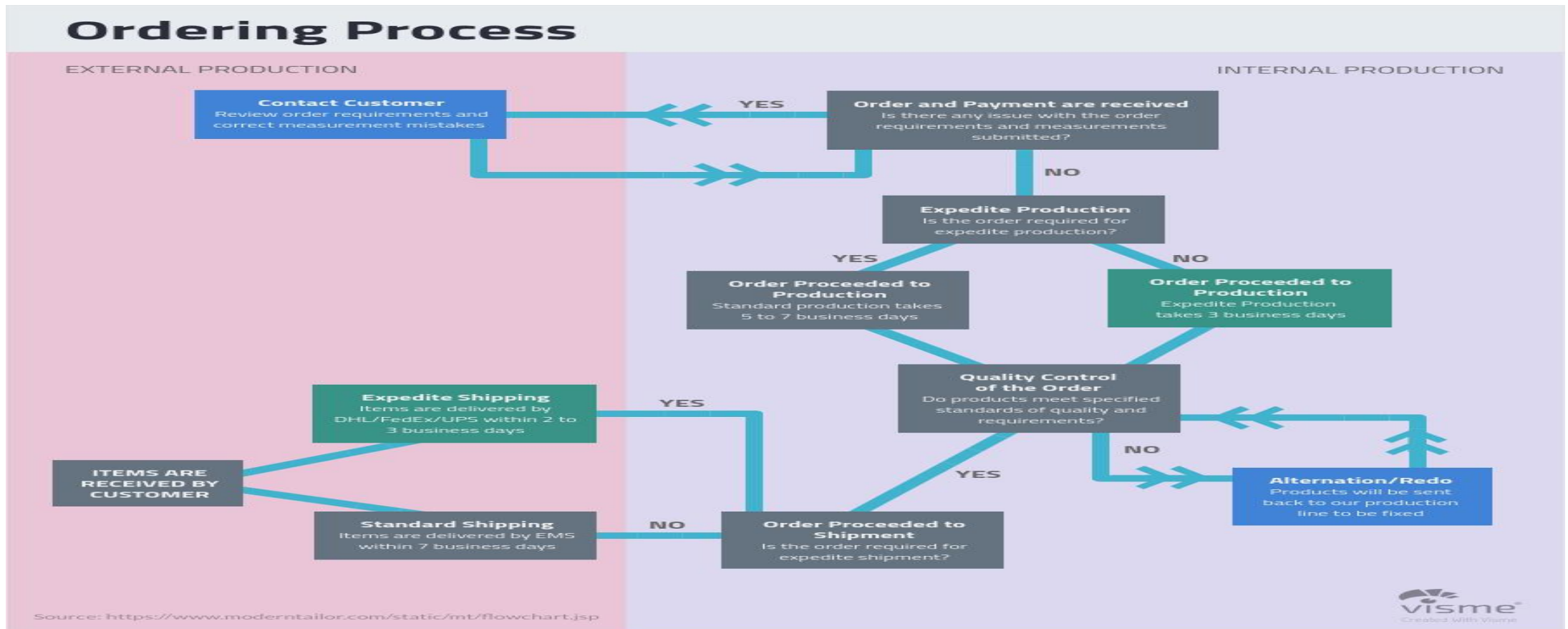
# Visualization tools

## What Are Data Visualization Tools?

Data visualization tools provide [data visualization designers](#) with an easier way to create visual representations of large data sets. When dealing with data sets that include hundreds of thousands or millions of data points, automating the process of creating a visualization, at least in part, makes a designer's job significantly easier.

These data visualizations can then be used for a variety of purposes: [dashboards](#), annual reports, sales and marketing materials, investor slide decks, and virtually anywhere else information needs to be interpreted immediately.

# Ordering process



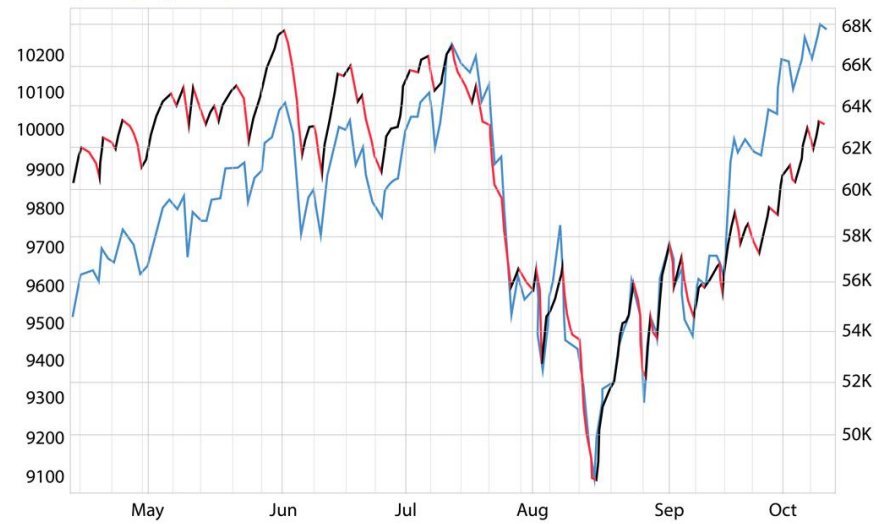
# Stock charts

**\$NYAD** (NYSE Advance-Decline Issues) INDX

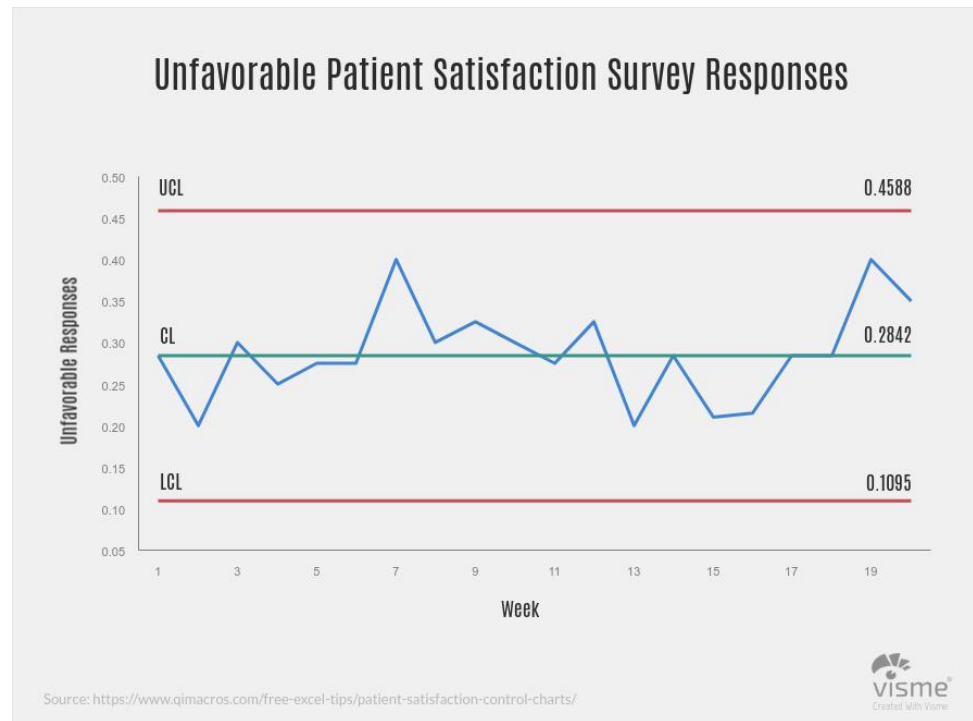
10-Oct-2007 **O** 63021.00 **H** 63021.00 **L** 63021.00 **C** 63021.00 **Chg** -145.00 (-0.23%) ▼

— **\$NYAD** (Daily) Cumulative 63021.00

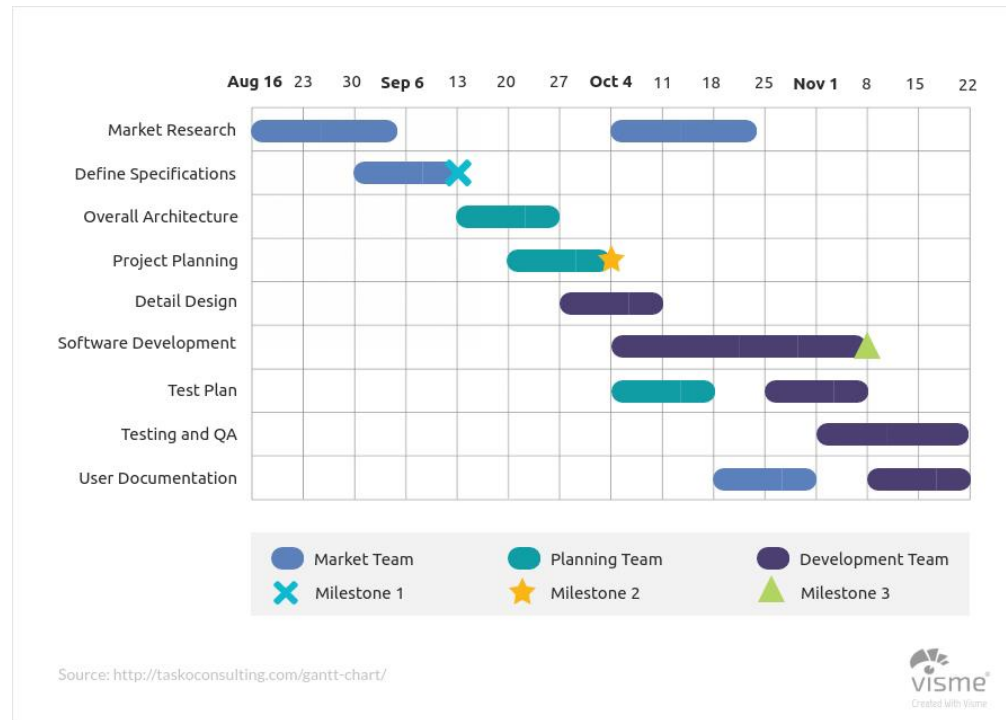
— **\$NYAD** (Daily) 10264,50



# Survey responses

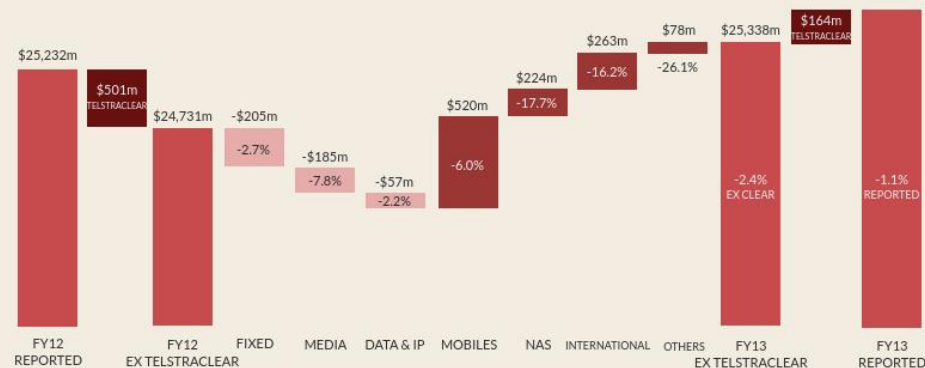


# Gantt charts



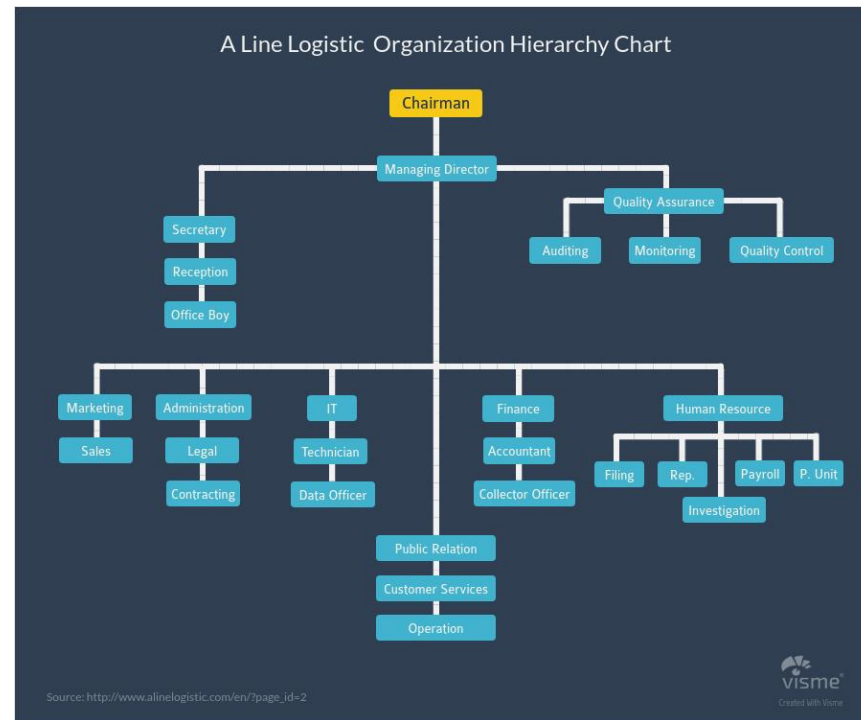
# Salers revenue growth

## PRODUCT PERFORMANCE SALES REVENUE GROWTH



Source: <http://www.fatprophets.com/Member%20Area/Product%20Landing/Report%20List/Report%20Page/Article%20Page.aspx?id=ccb67548-44d0-427e-b577-fc44ed90c32f&product=Australasian%20Equities&pt=paid>

# Hierarchy chart





# Background context

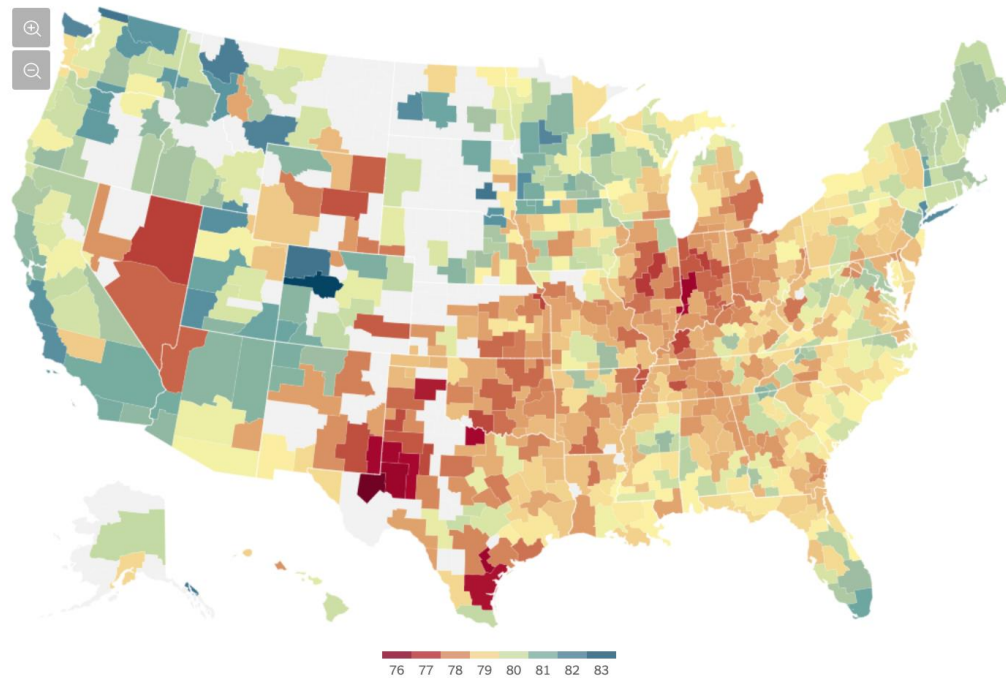
## 6.6.2 Background and Context

Context is something universally appreciated when being presented with any type of new information, and with data visualizations context is necessary for both basic understanding and proper interpretation. Context show as early as the title and hook of the article. Background is similar to context in how critical it is needed to avoid any poor interpretation, however it shows later on as a longer description.

Background begins with what drove the decision to use the subject matter. Even for purposes of work or schooling, the data used has a background which needs to be represented. Asking questions such as what history is behind this data, or where and from whom it is collected can provide the reader with the path to understanding the point of the project.

Background is both the history behind data or the history project creation itself, and context furthers the readers frame of mind. This frame of mind is necessary and is done such that they can understand how the data might be represented and how it can be interpreted it given the number of ways data is now provided.

# Background context



# Integration with business

## **6.6.4 Integration with Business**

(Lawton 2018)

### **6.6.4.1 Bring big data visualization up front**

Enterprises are finding ways to create data visualization front ends that can be explored by front-line workers. Rather than have developers create a purpose-built app, enterprises can wire up Google's Data Studio to their data sources, then create and style a simple UI with Analytics Canvas.

### **6.6.4.2 Connect time and space**

Marketers use location intelligence to understand consumer preferences, behavior or loyalty based on when, where and how often someone shows up. Customer support managers use location intelligence to create better customer experiences, since physical location is usually a big part of servicing a customer – from predicting arrival, delivering timely solutions to urgent issues and routing. Business development teams use location intelligence to reduce risk for future investments, such as where to open a new store, where to drill a new well, or where to construct a new cell tower.

### **6.6.4.3 Visualize the voice of the employee**