DA Bootcamp – Capstone 2

Released 2022-Mar-23

Background

InformatiTEK is a well-established market research firm that has developed a stable, dedicated base of clients across a wide range of organizations (Fortune 500 companies, government agencies, and non-profit). Their mission statement – "To boldly show what no company has shown before" – reflects their commitment to providing cutting edge insights to customers in a range of formats such as APIs and in-depth written analyses.

Recently, a gap seems to have surfaced in their offerings that puts that client base at risk. This comes at a time when change has been underfoot, with new leadership making a push for driving market share growth through rapid and aggressive salesmanship to prospective clients. That gap is on Data Visualization – InformatiTEK has traditionally provided static visualizations that work great on articles and PowerPoint slides, but lack the self-service appeal (as well as "wow" factor) of the interactive dashboards their competitors are offering.

To be sure, InformatiTEK certainly has in-house visualization specialists, but their skillsets are weighted toward D3.js. D3 is an open source (free) platform that is very developer oriented, enabling near limitless visualization possibilities, but with the trade-off of requiring deep investment in terms of time and man-hours to create those custom visuals. As a result, there's a fair amount of disconnect between the analysts (working with primarily with Excel) and the D3 developers.

InformatiTEK has released a Request for Proposal (RFP) seeking for a technology consulting partner that can not only align well to their culture, but also help drive change from within. To reach their client acquisition targets, they need to reduce the lead time to derive market insights for customers, as well as expose those insights to them. Therefore, **their ideal partner will not only provision a highly capable Data Analytics tech stack, but also prove they have analysts in-house that are resourceful, that know how to use the tools, and that could be quickly allocated to an ongoing Data Analysis program.**

TEKsystems has two leads on this client engagement – a Principal Consultant and a Solutions Architect. They have recruited a large team of TEK's finest data visualization consultants to help build a response to the RFP. The team's goal is to build a portfolio of "individually complete" data visualizations that will prove to the client that TEK is the right choice for InformatiTEK.

Requirements

This is a summary of the core functionality – for greater detail, check the Capstone 2 Rubric spreadsheet. When in doubt, ask.

As a consultant on the RFP Project, you're tasked with creating a data visualization proof of concept that leverages real-world raw datasets, to provide a compelling solution that showcases these essential features:

- Narrative: A clear motivation and a persuasive narrative through use of two or more Power BI Reports.
- Data visualization: high clarity and functionality. Well designed, purposeful, and easy to use. This is both for an analyst who might assume ownership in the future, and for those who might consume the dashboard.
- Statistical rigor: makes insightful claims, backed up with the use of one or more common statistical methods, such as correlation, regression, classification, ANOVA, etc.
- Data pre-processing: use of tools such as Pandas and Power Query to perform data transformations and quality checks.

You will present your solutions to the Project Leads and other TEKsystems leadership in a Q&A forum on Thursday March 31st and Friday April 1st. Use PowerPoint, then transition into Power BI and Python. It's strongly recommended you minimize the switching back and forth to PowerPoint. In the end, your dashboard will be collected into an online Visualization Gallery for submission to the client.

Your choice of topic should be guided by the following standards:

- Work appropriateness:
  - An industry with historical data
  - …or a social cause of importance to you
  - …or economic indicators
- From reputable sources ***(cited in your PowerPoint, your Dashboard, or both)***:
  - Data.World, Kaggle, etc **(though the true original source, if known, must be specified)**
  - Data from .gov's or other recognized authorities (US & Global)
  - *Mock or "toy" datasets are strictly not permitted*
- Data Visualization
  - Does the data have potential for visually intriguing dashboards?
  - What Visual Vocabulary techniques are possible? (See link in Resources)
- Statistical Merit
  - Does the data lend itself to statistical analysis?
  - Is there a story to tell? Are there surprising insights?
- Data Modeling/Data Clean-up requirements: What level of pre-processing is required, or possible?
  - Cross-domain viability: is it possible to combine two separate (but reputable) data sources? E.g. census data vs market data?

Stretch Functionality Requirement

Power BI has a wide range of features not covered specifically in the course, but which map directly to concepts covered during the bootcamp. You are required to branch out and incorporate some feature that is related to a topic from class, a Udemy lecture, or elsewhere. *For example, going deep on Power BI's DAX functions, using custom visualizations not found in base Power BI, Python integration, etc.* You must "claim credit" for it by dedicating a slide to it.

Submission

Submit all work, including Python Scripts, Jupyter Notebooks, Power BI files, and PowerPoint decks, to TEK Academy as a zip file. **Due date/time is 8:30am CT on Thursday, March 31st.**

Honor Code

We are committed to fairness. Your submission to TEK Academy is final, and no changes are permitted prior to your presentation.

Suggested Resources (Find Power BI equivalents)

Tableau

- Tableau vs D3: https://www.arkatechture.com/blog/tableau-vs.-d3
- Tableau Visual Vocabulary: https://public.tableau.com/en-us/gallery/visual-vocabulary
- Tableau Visual Best Practice: https://help.tableau.com/current/pro/desktop/en-us/visual_best_practices.htm
- Tableau filters: https://help.tableau.com/current/pro/desktop/en-us/order_of_operations.htm
- Tableau + SQL + Python Udemy course (possible stretch opportunities): https://teksystemsbootcamp.udemy.com/course/python-sql-tableau-integrating-python-sql-and-tableau/
- Tableau How-To's (Basic to Advanced): https://public.tableau.com/en-us/s/resources

Statistics:

- Jose Portilla's Probability & Statistics Course: https://teksystemsbootcamp.udemy.com/course/probability-and-statistics-for-business-and-data-science/learn/lecture/10741736#overview