Tableau Version Links

Version 1:
https://public.tableau.com/profile/courtney.ferguson.lee#!/vizhome/PISAVisualizations/PISAStory

Final Version:
https://public.tableau.com/profile/courtney.ferguson.lee#!/vizhome/PISAVisualizationsVersion3/ProgramforInternationalStudentAssessmentPISAExploration

Summary

This Tableau story visualizes student data from the 2012 PISA assessment. I cleaned the data in Python, including changing variables to human readable names (CNT → Country code 3-character), consolidating city / state data into country data (Florida (USA), Massachusetts (USA), United States of America → United States) and changing unrecognizable country codes into recognizable ones (Korea → South Korea, Chinese Taipei → Taiwan, etc.). The Tableau story starts with an overview of the results, then explores gender differences, student home life and student school life.

Design

In the first story point, I used a choropleth map to give readers an overview of each country's performance. Countries are colored using a brown to blue divergent gradient to make the data accessible to colorblind people. I also included a web frame so readers could learn more about each country's educational system by clicking on the map. Originally, there was a bar chart that displayed each country's score, but I felt it was redundant. I changed it to a plot of country rankings by subject after receiving feedback.

The choropleth of female vs male differences across subjects helped highlight gender differences in each country. This time I used pink to represent countries where females scored higher and blue to represent countries where males scored higher so it was easier to understand. I included a scatter chart in the bottom right to show how out of school study time affected males and females. Based on the feedback I received, I added a subheading letting the audience know that the size of each bubble represented the amount of time spent on the internet. I also relabeled the x-axis to let the audience know that homework study time was in total and not individual by subject.

In the student home life story point, I compared score percentiles to number of possessions to show global trends rather than country-specific trends. Originally, I compared the top 5 and bottom 5 countries, but changed it after receiving feedback. I used a line graph here for the survey data to show trends across each percentile. The boxplot was used to show how each survey question interacted with parent education levels. I only used two categories in the pie chart because I wanted the reader to be able to make quick, simple comparisons across each percentile.

The student school life story point was meant to show how classroom environments varied. I removed grid lines from the scatterplot and bar graphs to reduce chart clutter and stuck with the same color scheme used throughout the story for continuity.

Feedback

I received feedback from three friends. The first had trouble viewing the story on his phone, but once he switched to a desktop he did not have any issues. He did ask me what the sizes of the bubbles meant in the last dashboard and after realizing I did not make this explicit in the original visualization, I decided to edit it. The second and third friends offered several helpful suggestions:

Lee:

I like it. I noticed that China has the highest score. The one thing that was a little wierd was the graph started with all rows being dark blue but then they where blank until you clicked on them

looks like it's working. Its just very slow on my phone. It good on pc. Btw, what do the sizes of the bubbles mean on the last chart?

Sean:

Story Point 1

1. Performance levels for the countries remain pretty consistent cross all three disciplines, which discredits some of the stereotypes that Country A is particularly gifted in X but lacking in Y. Doesn't seem like certain countries are excelling in one area while flopping significantly in others.

Also, damn, China, slow down, time-out! We're tired over here! U.S. isn't performing abysmally, like the continual liberal narrative suggests, but we are certainly not ranking relative to the esteem that blindly-patriotic conservatives have about the U.S. being the greatest country. Anecdotal evidence does seem to suggest that we are in a worse position than these numbers suggest, though.

2. How was this data compiled? Has it be skewed in any way to account for capita? Was the test uniform across all countries, or tiered in its difficulty? If so, based on what factors?

3. The data suggests that the world powers/developed countries tended to attain higher levels of performance.

4. Based on the data interpretation from question 3, it seems as though a catch 22 exists, in which countries that are under-developed are so due to lack of widespread educational achievement levels,

and widespread educational achievement levels are difficult to attain because they are under-developed.

Story Point 2

1. The trending of the survey responses for each question remains very consistent across all three disciplines. As such, a different graphic representation that combines the three disciplines together and then shows the grouping of Score Range Groups Vs. Answer might be more impactful, with possibly filters on country? Example, students who scored above average were much more likely to claim to -not- know what Subjunctive Scaling, as they weren't in the practice of overclaiming, self-inflating their ego, and therefore better able to learn and absorb new (real) concepts. That isn't to say this graphic isn't impactful, there is some interesting data here, for sure. Perhaps the above would have been a neat flip-side graphic.

2. A couple of the items are a bit ambiguous. "Family Demands and Problems" with a Strongly Disagree to Strongly Disagree spectrum - does this mean a student who Strongly Disagrees is saying his/her family does not really have them, or that these demands/problems do exist but they do not significantly affect his/her ability to learn? I suppose a bit of context is lost. Also, not being a math nerd like you, I had to look up Subjunctive Scaling to learn that it was a red herring question, to see if students would overclaim knowledge.

3. Students that come from families that are more "successful" tend to perform better, via access to more resources. Students with better discipline (lesser truancy frequency, agree that trying hard is important) and humility (answered "Never heard of it" for Subjunctive Claiming) also tend to perform better. Finally, students with a more stable home-life tend to perform better.

4. While the graphic shows that the factors hit upon in the questions certainly were impactful in the students' scores, they were not smoking guns - students with higher scores were still represented in the groups of "negative" answers.

Story Point 3

1. Focusing particularly on the time spent studying out of school, it seemed that females spend more time studying, while achieving lower performance levels in Math and Science yet higher performance levels in Reading.

2. I am assuming these study-time hours were based on survey questions, with the students indicating their average hours spent in general, not per-discipline. Might want to specify that, it does seem that way based on the static grid clusters across the three disciplines. If that assumption is not correct, specify the correct context.

3. When societal context is applied, these findings make sense - the Math and Science findings make sense based on widespread cultural normatives of focusing on males for stimulating interest in the

maths/sciences (starting so early with marketing of toys) while the same applies to females in Reading, as there is a (diminishing, yet still prevalent) message of "kick up your feet and let the boys do the work."

4. The straight-forward message seems to indicate that males skew toward the maths/sciences while females skew toward the arts, agreeing with the stereotype, it is a raw-data message that doesn't take into account the factors (touched on in answer #1) that are causing the real-world to perpetuate that stereotype.

Marc:
1. What do you notice in the visualization?

This box loads on my browser (Chrome) right justified and makes me click and scroll over to start. I don't know what PISA is without having to read the font which is pretty small. I don't really know what I am clicking before I click it because I don't see any kind of header. It would also be hard to share verbally which box I wanted someone else to review without a header. I didn't even know the data was going to change once I clicked the boxes. The initial load time of the data was pretty long but was fast after cached. Having to click each box I like the color scheme and the format looks really good. The flow and the feel of everything looks very professional.

2. What questions do you have about the data?

There is plenty of information here for me to make an informed decision.

3. What relationships do you notice?

The relationship between books and TVs were interesting, it seems that having over 100 books matters a lot but the number of TVs matters little.

4. What do you think is the main takeaway from this visualization?

To give the audience a reasonable method to compare and contrast countries PISA scores.

5. Is there something you don't understand in the graphic?

No I think things were very clear

Resources

https://www.tableau.com/learn/training

http://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf

http://www.ipi.lu.lv/fileadmin/_migrated/content_uploads/Datu_analizes_rk.pdf

http://www.oecd.org/pisa/

http://mi2.mini.pw.edu.pl:8080/SmarterPoland/PISAcontest/