World Food consumption trends

Summary report prepared by Alina Bolat for Udacity Data Analytics Nanodegree.

Summary

World population is projected to grow over 11 billion by the end of the century [3]. The data has shown that there is a strong connection between food consumption and population growth, especially in China, India, and the USA. There was a noticeable drop in food availability in Post-soviet countries in the 90s. With the rapid increase in population, food consumption and its supply are becoming an increasingly important issue. **24%** of world's total greenhouse gas emissions are attributed to agriculture, forestry and food waste [4], thus contributing to global warming. Conscious regulation of diet choices and waste management should be a priority of every household in order for this issue not to have an astounding effect on future generations. The main inspiration of this project was to bring awareness and visually show the enormity of food and feed consumption rates.

Initial version 1 of tableau story can be found through the following link:
https://public.tableau.com/profile/alina.bolat#!/vizhome/Worldfoodconsumptiontrends1963-2013v1/Story1
Intermediate version 2 of tableau story can be found through the following link:
https://public.tableau.com/profile/alina.bolat#!/vizhome/World_food_consumption-DAND_final/Story1
Final version 3 of tableau story can be found through the following link:
https://public.tableau.com/profile/alina.bolat#!/vizhome/Worldfoodconsumptiontrends1963-2013v3/Howmuchfoodweconsume

Data

The Food and Agriculture Organization of the United Nations (FAO UN) has been gathering comprehensive data on food and agriculture for over 200 countries, starting from 1961 to 2013 (latest update) [2]. An extract from the Food Balance Sheets dataset is used for the visualisations [1]. It represents amount (weight in 1000 tons) of food/feed available for human/animal consumption per year for each country which is a member of the UN.

Initial dataset had to be cleaned on following aspects:

Tidiness:
- Year columns showing the quantity of food and feed consumption to be melted
- Unnessesary columns removed

Quality:
- 'Y' before the years to be removed
- `Item` values to be further grouped into categories, as current items are inaccurate and too numerous.

To elaborate more on food categories grouping: original food items had generic values such as 'Spices, Other' and 'Spices', along with more detailed ones like 'Pepper' and 'Cloves', so I've made a decision to group the food items into larger groups for ease of visualisations. The grouping was based on the nutrition summary guidelines from FAO UN website [5].
This process is well documented in the attached Jupyter notebook 'Data_wrangling_World_food_consumption.ipynb'.

Following reviews from other users I have also extracted additional information and joined it up with original dataset.
The wrangled and reshaped dataset consists of 7 variables and over 1 million observations. The food types have been further grouped into 13 categories and divided into further two groups: Food - supplies for human consumption and Feed - supplies for animal feed and production of other commodities. (For more information, please refer to the attached Jupyter notebook 'Data_wrangling_World_food_consumption.ipynb').

Design

The main specifications I have set for my story design were:
- Simplicity / coherency - make the visualisations clean, on white background, try to use as minimal features yet still maintain rich content.
- Use of dashboards - to provide user with controls and consolidate information on a single page.
- User centred approach - provide user with instructions on how to use the visualisations, include short summaries, but let the user explore information relevant to their needs.

The inspiration for the final layout was in most part taken from several artists in Tableau Gallery. I have also participated in several Tableau visualisation webinars which helped enrich my visualizations technically in versions 2 and 3.

Story visualisation choices:

Introduction: Beautiful bubble graph visualisation, to grasp interest of the viewer, was created in version 2 and improved in version 3. Only countries from the 'Large economies' set were assigned a special colour, the rest of the countries were assigned gradient colours of the 'Red-Golden' tableau palette in alphabetical order. Next to the visualisation there is a brief summary about following visualisations, background and origins of data, and main resources.
World outlook: Geographic data, is the main feature of chosen dataset, so the map visualisation with shading by total food weight seemed most appropriate. Users have control to hover over any country or group of countries to view summary statistics in the tooltip, additionally they can filter the visualisation by year using the slider at the top right. Map also works as a filter for the line chart and below bar graph which shows distribution of total food/feed consumed in weight and divided by various food groups. Users are able to hover over each element of visualisation for more insights. This dashboard received most positive feedback from 3 out of 4 reviewers, as being very informative.
Consumption growth: This is a minimalistic line graph illustrating growth among countries with biggest economies. There is a brief summary of most noticeable trends, the users can hiver over the line at any point for more insights. Initially this graph showed all countries in dataset, with control to choose any one country by user, however this was modified in latest version following the feedback.
Dairy consumption: This is a minimalist line graph, added in version 2. Alike the one before it shows growth trends among large economy countries in one of the most popular food groups - dairy. This shows a very interesting rapid growth of dairy products in India. Again there is a brief summary in the text box and users can see more insights by hovering over any point on the line.
Post-soviet world: This is an insight summary, I personally found most interesting, as Kazakhstan in my home country. The users are presented with an overview of food/feed supply in the 90s in Russia, Kazakhstan, Ukraine and Kyrgyzstan, a short background summary is provided at the top left, and lowest points in feed production are marked on the line charts on the right side of the page.
Food groups: This slide provides a breakdown of food/feed consumed by category. Initially the visualisation has bubble graph, however was changed to treemaps in version 3 following the review, it now better depicts proportions of food groups. Each section of treemap has % of total volume of consumables, and the user has control over filters and hovering for more insights. The bar graph is divided into food and feed categories, this allows for an easy overview when watching consumption per group.
Population growth: This section was added following a feedback from version 2, and also serves as a good conclusion page. This is large a map which allows user to hover over each country to see summary statistics, as well as observe the relationship between population growth and total consumption in the tooltip. A short summary is provided at the top of the page.

Feedback

There were two rounds of feedback gathering: version 1 review and version 2 review, final version is number 3. Overall reviewers liked the color choice, concise summaries and the use of dashboard to intuitively navigate and interpret the information. I received feedback from 4 sources:

Anuar Bolatov: award winning UX designer, v1 and v2 review. (verbal)

1. Make more informative legends.
2. Consider adding numerical summary caption inside the data visualisation, like figures of means, medians, sums, etc, in large fonts.
3. In Food groups: Add percentage into the bubble graphs, more information is required as to how much of each food type there is per country.
4. Change bubble graphs to something else. I find it difficult to view this chart and there is too much white space.
5. What about food consumption relative to the country population?

Asiya Zhunus: Medical student, v1 review. (verbal)

6. Make a single page infographic out of it, with pictures and illustrations.

Mohit Patel: DAND student v2 review. (Slack https://udacitydatascience.slack.com/archives/C72AP9J3Y/p1529925998000119)

7. On the consumption growth, it might be better if you select a couple of countries only including Russia. It makes Russia easy to find on the chart since you're going to talk about it later in the story.

Jaclyn Goddard: DAND student, v2 review. (Slack https://udacitydatascience.slack.com/archives/DBD6TUJ7K/p1529955747000300**)**

8. In the World Outlook page, I like how you also included instructions to scroll to see the whole dashboard. It makes it extremely user friendly.
9. Under Consumption Growth, I'd suggested changing how the Element filter is displayed. Maybe the dropdown version would take up less space without removing function. I think the different colors of the countries is the most powerful part of this message so that should definitely be open like you have it.
10. Under Dairy Consumption, I see minor typos in your text.

All feedback was very helpful. I have incorporated 8 out of 10 items from feedback, leaving out image enriched infographics and large summary captions, as this slowed down the story loading time and performance online.

Reflection

Overall the FAO dataset was my most favourite to work with. Not only was it relatively clean, and easy to wrangle, it was also of an area in which I am particularly interested. I found it very difficult to stop improving my visualisations, there is still a lot of potential in this dataset.

Data: After data wrangling phase was finished, I realised that JSON file format might have been better for this dataset, as, for instance, variables 'Year' and 'Population (1000 ppl)' are duplicated for each food item type, which is not ideal and takes up a lot of memory.
There is still so much more to explore in this data, for instance original FAO UN dataset consists of over 11 million entries, which include consumption per capita, larger geographic groups, import/export statistics, economy groups and more.
Future revisions of this project will include exploration into:

- What are food consumption trends per country capita?
- What portion of food is imported vs produces domestically?
- Which countries benefit from imports, and which countries are world's top domestic producers?
- What are the food consumption trends between economic groups?
- Communicate the compound annual growth rates (CAGR) and other summative statistics in an interesting way.
- Visualisations enriched with images and illustrations.

Feedback: I should have recorded the first round of feedback and posted it only, which I completely forgot to do.

Design: I have initially started making visualisations head on, without creating wireframes or sketching, which was a mistake as it took me much longer to figure out best visualisations and how to present them. I learned that planning and creating initial visualisations on paper allows for better quality of work and final visualizations.
It took a lot of Tableau worksheets and dashboards to make the final story, which does affect the loading time of the story online. For instance the 'Post-soviet world' dashboard took 5 worksheets to make as each country had its own graph. In the future will aim to visualise several graphs in a single worksheet and avoid hard coding the visualisations.

Resources

[1] The initial dataset provided by Dor Oppenheim
https://www.kaggle.com/dorbicycle/world-foodfeed-production
[2] More information: Food and Agriculture Organization of the United Nations
http://www.fao.org/faostat/en/#home
[3] Max Roser and Esteban Ortiz-Ospina (2018) - "World Population Growth". Published online at OourWorldInData.org. Retrieved from:
https://ourworldindata.org/world-population-growth
[4] Greenhouse Gas emissions: United States Environmental Protection Agency:
https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data
[5] Food Groups have been divided based on Food and Agriculture Organization of the United Nations:
http://www.fao.org/nutrition/education/food-dietary-guidelines/en/