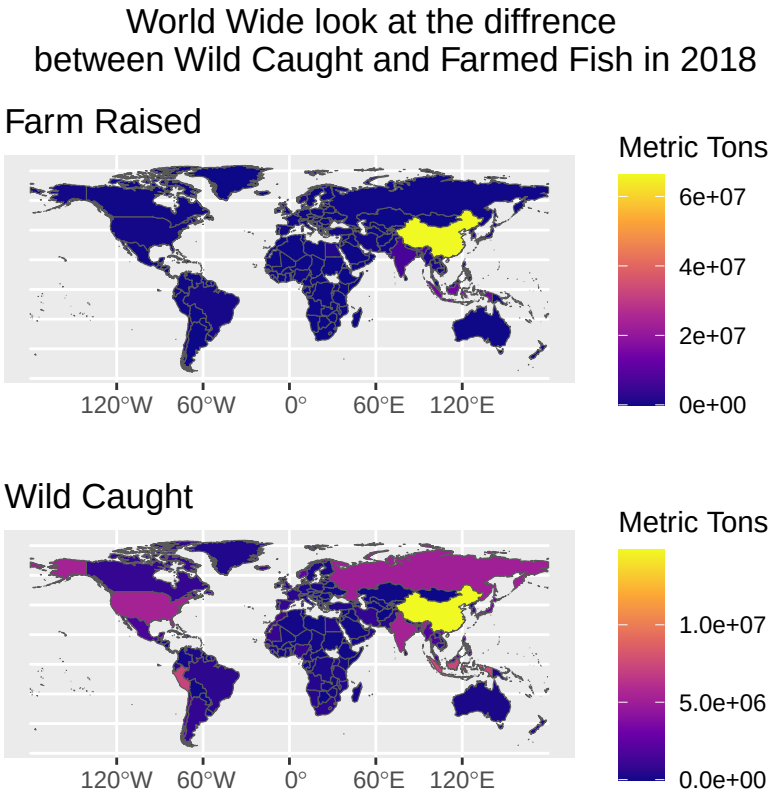
MXB262 Project 2024

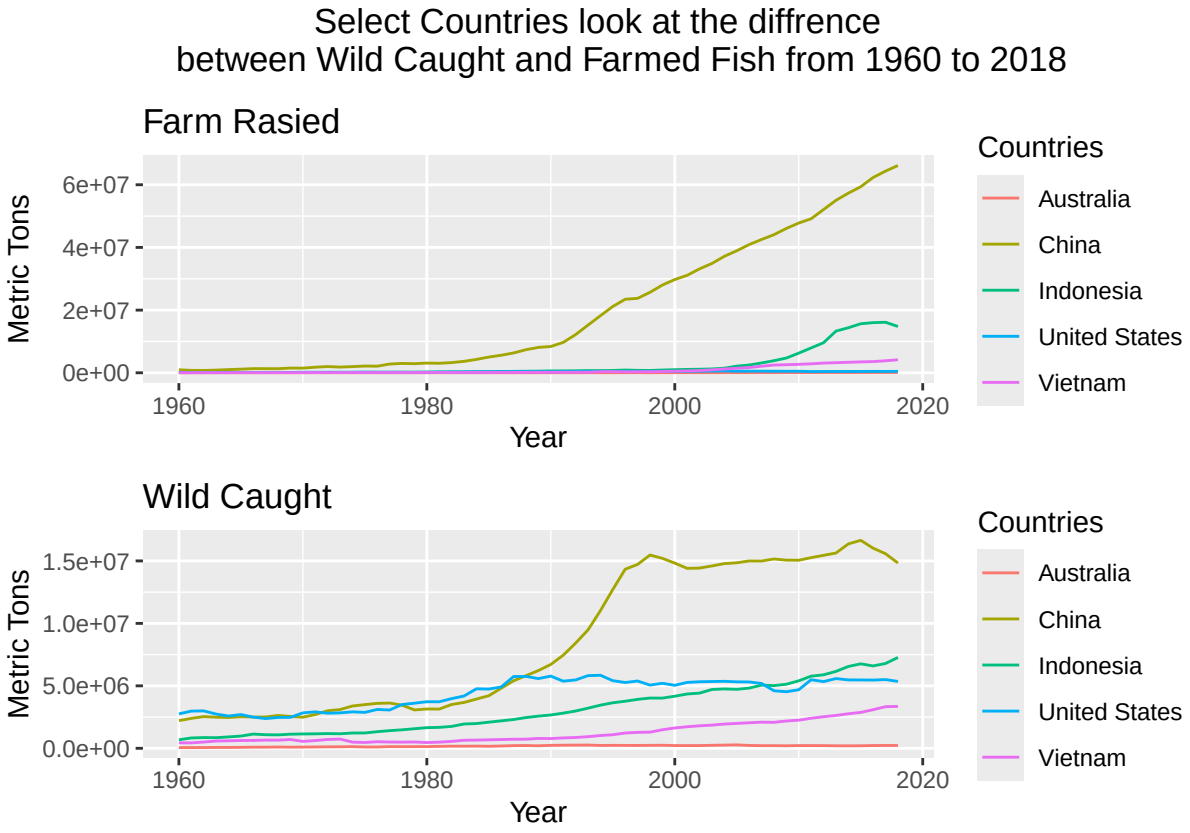
Braydan Newman - n11272031

```r
# Get Data
captured_vs_farmed <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/mas

# Rename Columns
colnames(captured_vs_farmed)[which(names(captured_vs_farmed) == "Aquaculture production (metric tons)")]
colnames(captured_vs_farmed)[which(names(captured_vs_farmed) == "Capture fisheries production (metric to

# Set NA as 0
captured_vs_farmed$AquacultureProduction[is.na(captured_vs_farmed$AquacultureProduction)] <- 0
captured_vs_farmed$CaptureProduction[is.na(captured_vs_farmed$CaptureProduction)] <- 0

# Create a subset of only 2018 data
data_2018 <-  captured_vs_farmed %>% filter(Year == 2018)
data_2018$total <- with(data_2018, AquacultureProduction - CaptureProduction)

# Create a subset of only select countries
data_countries <-  captured_vs_farmed %>%
  filter(Entity %in% c("Australia", "China", "Indonesia", "Vietnam", "United States"))
```

```r
# Graph 1 - "World view (2018)"
# Combine with world data to draw with countries
world <- ne_countries(scale = "medium", returnclass = "sf")
data_2018_world <- inner_join(world, data_2018, by = c("adm0_a3" = "Code"))

# Create farm raised world map
grown_fish_world <- ggplot(data = data_2018_world) +
  geom_sf(aes(fill = AquacultureProduction)) +
  scale_fill_viridis_c(option = "plasma") +
  labs(x =  "", y = "", fill = "Metric Tons",
       title = "Farm Raised")

# Create wild caught world map
capture_fish_world <- ggplot(data = data_2018_world) +
  geom_sf(aes(fill = CaptureProduction)) +
  scale_fill_viridis_c(option = "plasma") +
  labs(x =  "", y = "", fill = "Metric Tons",
       title = "Wild Caught")

# Combine both maps together
grown_fish_world / capture_fish_world +
  plot_annotation(
    theme = theme(plot.title = element_text(hjust = 0.5)),
    title = 'World Wide look at the diffrence
    between Wild Caught and Farmed Fish in 2018',
  )
```




```
# Graph 2 - "Select Countries Trends"
# Create line graph for both "Wild Caught" and  "Farm Raised"
grown_fish <- ggplot() +
  geom_line(data_countries,
            mapping = aes(x = Year, y = AquacultureProduction, color = Entity)) +
  labs(x =  "Year", y = "Metric Tons", title = "Farm Rasied", color = "Countries")

capture_fish <- ggplot() +
  geom_line(data_countries,
            mapping = aes(x = Year, y = CaptureProduction, color = Entity)) +
  labs(x =  "Year", y = "Metric Tons", title = "Wild Caught", color = "Countries")

grown_fish / capture_fish +
  plot_annotation(
    theme = theme(plot.title = element_text(hjust = 0.5)),
    title = 'Select Countries look at the diffrence
    between Wild Caught and Farmed Fish from 1960 to 2018',
  )
```

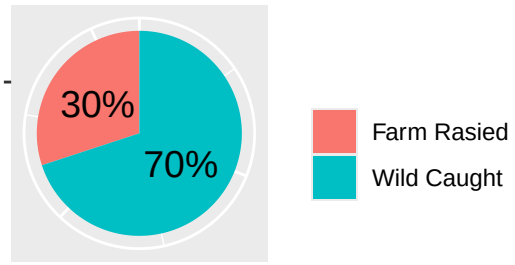



```r
# Graph 3
# Create subset of data for each of the select countries
Australia_single = data_2018 %>% filter(Entity == "Australia")
China_single = data_2018 %>% filter(Entity == "China")
Indonesia_single = data_2018 %>% filter(Entity == "Indonesia")
Vietnam_single = data_2018 %>% filter(Entity == "Vietnam")
United_States_single = data_2018 %>% filter(Entity == "United States")

# Turn the subsets into data frames with a different format
Australia_df <- data.frame(
  group = c("Farm Rasied", "Wild Caught"),
  value = c(Australia_single$AquacultureProduction, Australia_single$CaptureProduction)
)
China_df <- data.frame(
  group = c("Farm Rasied", "Wild Caught"),
  value = c(China_single$AquacultureProduction, China_single$CaptureProduction)
)
Indonesia_df <- data.frame(
  group = c("Farm Rasied", "Wild Caught"),
  value = c(Indonesia_single$AquacultureProduction, Indonesia_single$CaptureProduction)
)
Vietnam_df <- data.frame(
  group = c("Farm Rasied", "Wild Caught"),
  value = c(Vietnam_single$AquacultureProduction, Vietnam_single$CaptureProduction)
)
United_States_df <- data.frame(
  group = c("Farm Rasied", "Wild Caught"),
  value = c(United_States_single$AquacultureProduction, United_States_single$CaptureProduction)
```

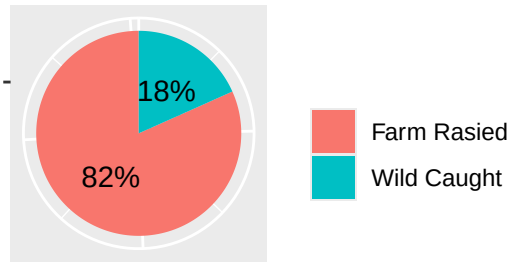
```
)

# Create a pie graph for each country
Australia_plot <- ggplot(Australia_df, aes(x="", y=value, fill=group)) +
  geom_bar(stat="identity", width=1) +
  coord_polar("y", start=0) +
  theme(axis.text.x=element_blank()) +
  geom_text(position = position_stack(vjust = 0.5),
            aes(label = percent(value/sum(value))), size=5) +
  labs(x =  "", y = "", title = "Australia 2018", fill = "")

China_plot <- ggplot(China_df, aes(x="", y=value, fill=group)) +
  geom_bar(stat="identity", width=1) +
  coord_polar("y", start=0) +
  theme(axis.text.x=element_blank()) +
  geom_text(position = position_stack(vjust = 0.5),
            aes(label = percent(value/sum(value))), size=4) +
  labs(x =  "", y = "", title = "China 2018", fill = "")

Indonesia_plot <- ggplot(Indonesia_df, aes(x="", y=value, fill=group)) +
  geom_bar(stat="identity", width=1) +
  coord_polar("y", start=0) +
  theme(axis.text.x=element_blank()) +
  geom_text(position = position_stack(vjust = 0.5),
            aes(label = percent(value/sum(value))), size=5) +
  labs(x =  "", y = "", title = "Indonesia 2018", fill = "")

Vietnam_plot <- ggplot(Vietnam_df, aes(x="", y=value, fill=group)) +
  geom_bar(stat="identity", width=1) +
  coord_polar("y", start=0) +
  theme(axis.text.x=element_blank()) +
  geom_text(position = position_stack(vjust = 0.5),
            aes(label = percent(value/sum(value))), size=5) +
  labs(x =  "", y = "", title = "Vietnam 2018", fill = "")

Australia_plot + China_plot + Indonesia_plot + Vietnam_plot +
  plot_annotation(
    theme = theme(plot.title = element_text(hjust = 0.5)),
    title = 'Select Countries look at the Percentages
    between Wild Caught and Farmed Fish in 2018',
  )
```

Select Countries look at the Percentages
between Wild Caught and Farmed Fish in 2018

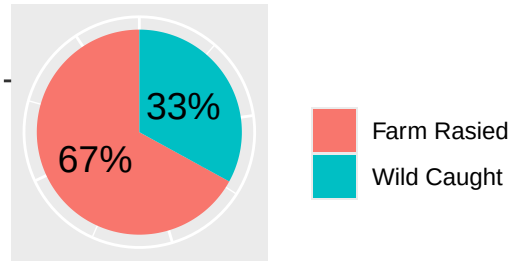
Australia 2018



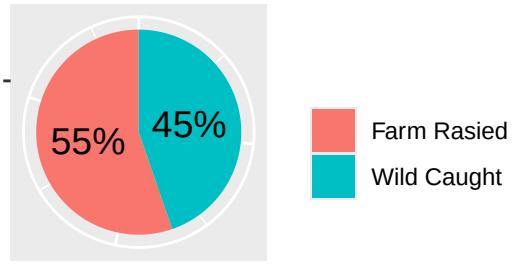
China 2018



Indonesia 2018



Vietnam 2018



Main Message and aim of Communication

The main message communicated through these visualizations is the significant and growing role of aquaculture (farm-raised fish) in global fish production, compared to traditional wild-caught fishing.

The visualizations clearly communicate that while traditional wild-caught fishing remains important, there is a substantial and rapidly growing shift towards aquaculture, especially in countries like China. This shift is driven by the increasing global demand for fish and the scalability of farm-raised fish production. Understanding these trends is crucial for stakeholders in the fishing and aquaculture industries, as well as policymakers focused on sustainable fishery practices and food security.

These insights emphasize the need for continued investment and innovation in aquaculture to meet future fish demand sustainably, while also addressing the environmental and economic impacts of this shift.

Description of the dataset

Source: https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2021/2021-10-12/capture-fisheries-vs-aquaculture.csv

The dataset captures the production of aquaculture and capture fisheries across various countries and years. Some columns have missing values, particularly the "Code" and "Aquaculture production" columns.

Columns:

- **Entity**: The country or region name.
- **Code**: A three-letter country code (ISO 3166-1 alpha-3)
- **Year**: The year of the data entry.
- **Aquaculture production (metric tons)**: The production of fish and other aquatic life under controlled conditions for that year, measured in metric tons.
- **Capture fisheries production (metric tons)**: The production from capturing wild fish and other aquatic life, measured in metric tons.

Who is the intended audience?

The intended audience for these visualizations includes policymakers and government officials, environmental and conservation organizations, researchers and academics, industry stakeholders in fisheries and aquaculture, and the public and consumers. Policymakers and government officials need data-driven insights to craft effective policies that balance economic growth with environmental conservation. Environmental and conservation organizations focus on promoting sustainable practices and protecting marine life. Researchers and academics provide foundational research that informs policies and industry practices. Industry stakeholders are directly impacted by trends and need to adapt strategies to remain competitive and sustainable. The public and consumers drive demand through their awareness and choices, influencing both policy and industry practices. By targeting these audiences, the visualizations aim to provide a comprehensive overview of current fish production trends, encouraging informed decision-making and fostering a collaborative approach to sustainable fishery management.

What is the intent behind these visualizations?

The intent behind these visualizations is to highlight the significant differences and trends in global fish production between farm-raised and wild-caught methods. By presenting a clear comparison through maps, time-series data, and pie charts, the visualizations aim to inform policymakers, environmental organizations, researchers, industry stakeholders, and the public about the growing importance of aquaculture, particularly in countries like China. They emphasize the rapid increase in farm-raised fish production and its dominance in certain regions, contrasting it with the more stable growth of wild-caught fish. This information is crucial

for understanding the implications on sustainability, food security, and economic strategies within the fishery sector, thereby supporting informed decision-making and fostering a collaborative approach to managing fish resources sustainably.

How are these visualizations a good communication tool?

These visualizations are an effective communication tool because they combine clear, visually appealing aesthetics with robust, insightful data. The use of color gradients in the world maps effectively distinguishes between different levels of fish production, making it easy to identify regions with high and low outputs immediately. The time-series charts for selected countries provide a historical perspective, showing trends over nearly six decades and highlighting the rapid rise of aquaculture, particularly in China. This long-term view is crucial for understanding how fish production methods have evolved. The pie charts offer a straightforward comparison of farm-raised versus wild-caught fish production for specific countries, using contrasting colors to emphasize differences clearly. These visual elements are carefully chosen to ensure that the data is accessible and understandable, even for those without a deep background in fisheries or aquaculture. Moreover, the combination of global and country-specific data allows for a comprehensive analysis, catering to a wide range of audiences from policymakers to the public. By presenting complex data in a clear and visually engaging manner, these visualizations facilitate informed decision-making and promote a better understanding of global fish production trends.

How should these visualizations be interpreted?

These visualizations should be interpreted as a comprehensive comparison of global fish production, highlighting the differences between farm-raised and wild-caught fish. The world maps show the distribution of fish production in 2018, with China leading significantly in both categories but with a much higher output in farm-raised fish. The time-series charts illustrate the historical trends from 1960 to 2018, indicating a steep rise in farm-raised fish production, especially in China, contrasted with a steadier growth in wild-caught fish production. The pie charts provide a snapshot of the relative contributions of farm-raised and wild-caught fish in 2018 for selected countries, showing varying degrees of reliance on each method. For instance, China relies heavily on farm-raised fish, while Australia and Indonesia depend more on wild-caught fish. These visualizations collectively underscore the growing importance of aquaculture in meeting global fish demand and the diverse production strategies of different countries.

Justification of plot type and pre-attentive attributes you have chosen to communicate your message to your audience

Map plot - World Maps (2018):

The choice to use 2 map plots with the countries fill color representing the amount in tons of Farm Raised or Wild Caught fish, was because it gives the best way to show generally every countries levels without being crowded like a bar graph would be with every country on it, but it gives enough detail to show where the hotspots are. The world maps were chosen to visualize the global distribution of farm-raised and wild-caught fish production in 2018 because they provide a comprehensive geographic overview. The use of color gradients is a pre-attentive attribute that allows viewers to quickly discern areas of high and low production. Yellow indicates the highest production levels, while dark blue represents the lowest, enabling immediate visual differentiation. This map format effectively communicates the dominance of certain regions, like China, in both farm-raised and wild-caught fish production, making it easy for the audience to grasp spatial patterns and regional contributions immediately. The split maps were a way of showing both metrics I wanted to show. Farm-Raised and Wild Caught. I experimented with having it on the same map, but it became unusable and hard to read any data.

Time-Series Line Charts (1960-2018) :

The choice of the 2-line plots was crucial in showing in greater detail countries in comparison to each other, this was something that you can't get from the Map, just like the time aspect and how each country changes over time. It also illustrates trends in fish production from 1960 to 2018. However, something you can't do in a line plot is show every country and keep the graph readable, like you can in the maps. The countries chosen are ones that try to represent massive contributors to this market but also countries that would have a big reliance on fish as a resource of food. The choice of a line plot specificity was mainly because of the time axis needed. The Time aspect was needed in these visualizations, as to understand an industry, one needs to understand historically what that industry did. The best choice for looking at time and trends is a line graph, especially with multiple targets to plot. Each country is represented by a distinct color, leveraging pre-attentive attributes to differentiate between multiple data series easily. This approach allows the audience to track the progress of each country's fish production, observe the rapid growth in aquaculture, particularly in China, and compare it with more stable trends in wild-caught fish production. The visual clarity and historical context provided by these charts help convey the message effectively.

Pie Charts (2018):

The pie charts were selected to present the relative contributions of farm-raised and wild-caught fish for selected countries in 2018 and it is essential to sum up the previous 2 graphs. While looking at the pervious 2 graphs one might assume that China's Farm-Rasied and Wild Caught are ruffly the same however one is significantly higher, this is due to scaling of the axis.

Pie charts are effective for showing part-to-whole relationships, making them appropriate for illustrating the proportion of each type of fish production within a country. The use of contrasting colors—red for farm-raised and blue for wild-caught fish—serves as a pre-attentive attribute, allowing viewers to quickly understand the distribution. This visual format makes it easy to compare the reliance on different production methods across countries like Australia, China, Indonesia, and Vietnam. The simplicity and immediacy of pie charts ensure that the audience can quickly grasp the comparative data without needing extensive analysis. After looking at the Pie plots, you can understand that China had 4 times the quantity of Farm-Raised fish then Caught. Thats why the pie graphs are needed and complement the rest of the visualizations well.

Roadblocks

A roadblock I found was with quick interpretations of the graphs, specifically the scaling/proportion issue. When looking at the first 2 graphs understanding the difference in the amount of each of the quantities (Farm-Raised and Wild Caught) it is easy to miss that the scaling is not 1:1, without looking at the graph legend or the axis numberings between the 2 stacked plots. With the line plot China looks like it has the same quantity of Farm-Raised to Wild Caught but this is not the case, China has 4 time more Farm-Raised, and this is hard to see in the first two graphs. A solution to this could have been having the same scaling for each plot in the stacked visualization. After experimenting with this idea, I was unable to stop it, reducing the useability of the graph. With the same scale it stopped the wild caught plots from showing much difference in data points. The solution was to have the pie plots to show proportionality instead, make the all the visualizations work together.

Recommendations

First Recommendation:

More data, an end goal for these visualizations would be showing how destructive to nature Wild Catching of fish with Drag nets and overfishing can be. With more space and more visualization, a whole extra set of graphs showing the quantities of Wild Caught fish from different commercial fishing the techniques like dragnets and how overfishing has affected ecosystems. This would be a great addition to extend the current data and start showing what commercial fishing is doing and how it needs to be more regulated and where and who is most to blame for the loss of habitat.

Second Recommendation:

Predictions, for the line graphs, a great addition would be prediction on the feature, but not just a simple extrapolation of the data, but working out how proposed bills and regulations in different countries will affect these feature trends. This would be valuable information on what bills and regulations to vote for and what might make a difference.