InfoVis Final Project Data Exploration • Plastic Pollution

Kylie Lin, Sherry Xue, Anika Nerella, Nawar Alhaddad, Mackenzie Wood

Contents

Introduction

Our Final Project is guided by four questions: Where does polluting plastic come from? How much polluting plastic is there in the world and how much has there been over time? Where does the plastic end up eventually? Why should I care; i.e. where does the plastic go in my life? Each of the explored datasets below answers one of these guiding questions (two datasets are used to answer where plastic goes). **A final overview of trends of focus for our project can be found in the conclusion section of this PDF.**

Plastics by Industrial Sector + Polymer Types

Show 10 entries Search:

	entity	year	primary_plastic_production_million_tonnes
1	Additives	2015	25000000
2	All industrial sectors	2015	407000000
3	Building and Construction	2015	65000000
4	Consumer & Institutional Products	2015	42000000
5	Electrical/Electronic	2015	18000000
6	HDPE	2015	52000000
7	Industrial Machinery	2015	3000000
8	LD, LDPE	2015	64000000
9	Other polymer type	2015	16000000
10	Other sectors	2015	47000000

Showing 1 to 10 of 19 entries Previous 1 2 Next

Polymer Types

The data on polymer types is documented in 2015 and it shares the dataset with the plastic consumption by industrial sectors. It includes 2 variables: the entity (type of plastic: PET, HDPE, LDPE, PET etc.) and the amount of primary plastic production in million metric tonnes. Looking at only the data for polymers, the minimum primary plastic production is 6.80e+07 million tonnes (PP) and the maximum primary plastic production is 2.50e+07 million tonnes (PS). We plan on visualizing the polymer data with the every day products to provide users a more tangible understanding of the amount of plastic production by polymer types. The figures below are several ways that we've been exploring how to visualize the data through actual products.

Figure 1: PET visualization 1 with single use water bottle. This image illustrates one possible way to visualize the PET data. I got the number of plastic bottles by converting the data in the table to grams and then divided by the average weight of an 16oz plastic water bottle, which is 19 grams.

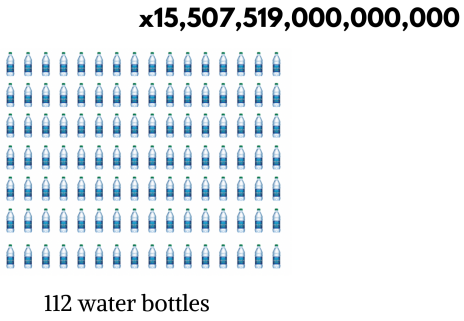


Figure 2: PET visualization 2. This figure shows an alternative way of illustrating PET data. It shows plastic water bottle in bulk as a single unit and then quantified the bulk.

Note: the visualization didn't take into account of the complexity of the production process as data on the exact number of plastic polymer used in plastic product production process isn't publicly available. Therefore, the visualization a rough estimate of the translation using the weight of plastic product (ex. Single- use plastic water bottle etc.) but didn't take into account for the other materials (ex. Oil and water) used in production.

For next steps, we plan to explore ways to visualize the data even more using metaphors such as the distance from earth to moon or the size of a continent to describe the numbers associated with each plastic product.

Figure 3: PP visualization with reusable plastic containers. This figure explores visualizing PP polymer data in reusable containers.

Figure 4: HDPE visualization with milk jar.

Figure 5: LDPE visualization with plastic trash bags.

```
skim_without_charts(sector_polymer)
```

Show 10 entries Search:

	Missingness	Industry/Polymer Type Count	Year	Production Range
1	No Missingness	19	Only 2015	3e+06 to 4.07e+08

Showing 1 to 1 of 1 entries Previous 1 Next

Industrial Sectors

The data from 2015 on plastic production by industrial sector gave some indication of which industries are the biggest producers of plastic. The data includes 2 variables: entity (industry/sector: packaging, textiles, transportation, building & construction, etc...) and the amount of plastic produced in million tonnes. This data was in the same table as the polymer types data but the desired variables were separated and an initial exploratory visualization using a bubble chart was preformed. As expected, packaging was the biggest industrial producer with 146 million tonnes. The smallest producer was the industrial machinery sector with 3 million tonnes of plastic.

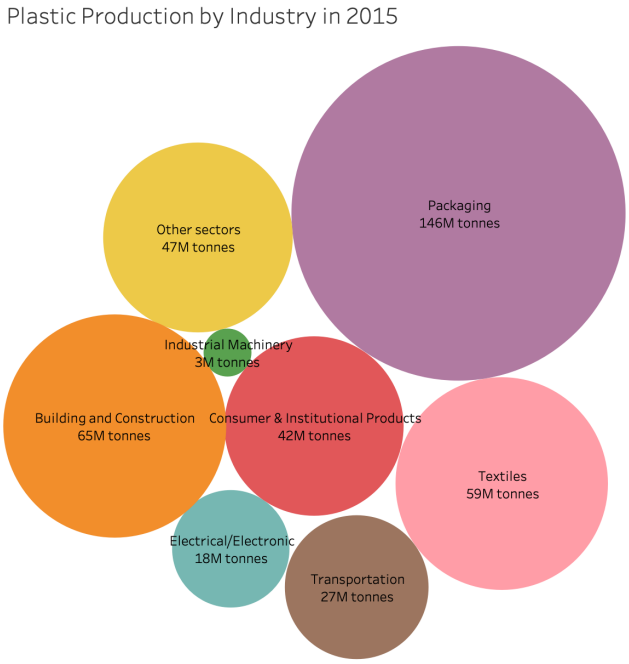


Figure 6: Plastic Production by Sector

Although this gives some idea of where plastic is coming from, it could be useful to couple this information with the polymer types data and try a stacked bar plot.

Production Over Time

Summary of Datasets

```
#Summarize datasets.
summary(plastics_type)
```

```
##     entity              code                year
##  Length:66          Length:66          Min.   :1950
##  Class :character   Class :character   1st Qu.:1966
##  Mode  :character   Mode  :character   Median :1982
##                                        Mean   :1982
##                                        3rd Qu.:1999
##                                        Max.   :2015
##  global_plastics_production_million_tonnes
##  Min.   :  2000000
##  1st Qu.: 20750000
##  Median : 76500000
##  Mean   :118530303
##  3rd Qu.:198500000
##  Max.   :381000000
```

```
summary(population_by_time)
```

```
##      loc_id         location             var_id         variant
##  Min.   :   4.0   Length:280932      Min.   :  2.00   Length:280932
##  1st Qu.: 300.0   Class :character   1st Qu.:  2.00   Class :character
##  Median : 586.0   Mode  :character   Median :  5.00   Mode  :character
##  Mean   : 703.1                      Mean   : 22.41
##  3rd Qu.: 903.0                      3rd Qu.:  9.00
##  Max.   :5501.0                      Max.   :207.00
##
##       time        mid_period      pop_male          pop_female
##  Min.   :1950   Min.   :1950   Min.   :       7   Min.   :       7
##  1st Qu.:2031   1st Qu.:2032   1st Qu.:    1848   1st Qu.:    1861
##  Median :2055   Median :2056   Median :   11368   Median :   11646
##  Mean   :2051   Mean   :2052   Mean   :  232436   Mean   :  229875
##  3rd Qu.:2078   3rd Qu.:2078   3rd Qu.:   85264   3rd Qu.:   86489
##  Max.   :2100   Max.   :2100   Max.   :10920004   Max.   :10712733
##                                NA's   :30056      NA's   :30056
##    pop_total         pop_density
##  Min.   :       0   Min.   :    0.05
##  1st Qu.:    1240   1st Qu.:   34.21
##  Median :   14206   Median :   91.59
##  Mean   :  412854   Mean   :  426.66
##  3rd Qu.:  117166   3rd Qu.:  216.99
##  Max.   :21632737   Max.   :56025.84
##
```

```
summary(waste2010)
```

```
##     entity              code                year
##  Length:168         Length:168         Min.   :2010
##  Class :character   Class :character   1st Qu.:2010
##  Mode  :character   Mode  :character   Median :2010
##                                        Mean   :2010
##                                        3rd Qu.:2010
##                                        Max.   :2010
##  plastic_waste_generation_tonnes_total
##  Min.   :     527
```

```
##  1st Qu.:   32512
##  Median :  283357
##  Mean   : 1626619
##  3rd Qu.: 1083424
##  Max.   :59079741
```

```
#Look at top of datasets for adjusting naming convention.
head(plastics_type) %>%
  datatable()
```

Show 10 entries Search:

	entity	code	year	global_plastics_production_million_tonnes
1	World	OWID_WRL	1950	2000000
2	World	OWID_WRL	1951	2000000
3	World	OWID_WRL	1952	2000000
4	World	OWID_WRL	1953	3000000
5	World	OWID_WRL	1954	3000000
6	World	OWID_WRL	1955	4000000

Showing 1 to 6 of 6 entries Previous 1 Next

```
head(population_by_time) %>%
  datatable()
```

Show 10 entries Search:

	loc_id	location	var_id	variant	time	mid_period	pop_male	pop_female	pop_total	pop_density
1	4	Afghanistan	2	Medium	1950	1950.5	4099.243	3652.874	7752.117	11.874
2	4	Afghanistan	2	Medium	1951	1951.5	4134.756	3705.395	7840.151	12.009
3	4	Afghanistan	2	Medium	1952	1952.5	4174.45	3761.546	7935.996	12.156
4	4	Afghanistan	2	Medium	1953	1953.5	4218.336	3821.348	8039.684	12.315
5	4	Afghanistan	2	Medium	1954	1954.5	4266.484	3884.832	8151.316	12.486
6	4	Afghanistan	2	Medium	1955	1955.5	4318.945	3952.047	8270.992	12.669

Showing 1 to 6 of 6 entries Previous 1 Next

```
head(waste2010) %>%
  datatable()
```

Show 10 entries Search:

	entity	code	year	plastic_waste_generation_tonnes_total
1	Albania	ALB	2010	73364
2	Algeria	DZA	2010	1898343
3	Angola	AGO	2010	528843
4	Antigua and Barbuda	ATG	2010	22804
5	Argentina	ARG	2010	2753550
6	Aruba	ABW	2010	9352

Showing 1 to 6 of 6 entries Previous 1 Next

Running Regressions on Data

```
##
## Call:
## lm(formula = plastic_waste_generation_tonnes_total ~ pop_total,
##     data = waste_bycountry2010)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -28219279   -446868   -409519   -101164  22850872
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4.491e+05  2.795e+05   1.607     0.11
## pop_total   2.614e+01  1.779e+00  14.696   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3313000 on 147 degrees of freedom
## Multiple R-squared:  0.595,  Adjusted R-squared:  0.5922
## F-statistic:   216 on 1 and 147 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = plastic_waste_generation_tonnes_total ~ pop_male +
##     pop_female, data = waste_bycountry2010)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -14427831   -146150    338324    429016  16576118
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -414094.3   246088.0  -1.683   0.0948 .
## pop_male      -1162.8      112.8 -10.305   <2e-16 ***
## pop_female     1293.7      120.3  10.753   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2580000 on 131 degrees of freedom
##   (15 observations deleted due to missingness)
## Multiple R-squared:  0.7794, Adjusted R-squared:  0.776
## F-statistic: 231.4 on 2 and 131 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = plastic_waste_generation_tonnes_total ~ pop_density,
##     data = waste_bycountry2010)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1454281 -1398920 -1189953  -512955 57632920
##
```

```
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1458477.75  434727.54   3.355  0.00101 **
## pop_density     -79.95     207.88  -0.385  0.70110
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5204000 on 147 degrees of freedom
## Multiple R-squared:  0.001005,   Adjusted R-squared:  -0.005791
## F-statistic: 0.1479 on 1 and 147 DF,  p-value: 0.7011
```

```
# Lineplot of total plastics production over time.
ggplot(plastics_time, aes(year, global_plastics_production_million_tonnes)) +
  geom_line() +
  labs(title= "Plastic Production by Year", x= "Year", y=" Plastic Production (in tonnes)") +
  scale_x_continuous(n.breaks=10)
```

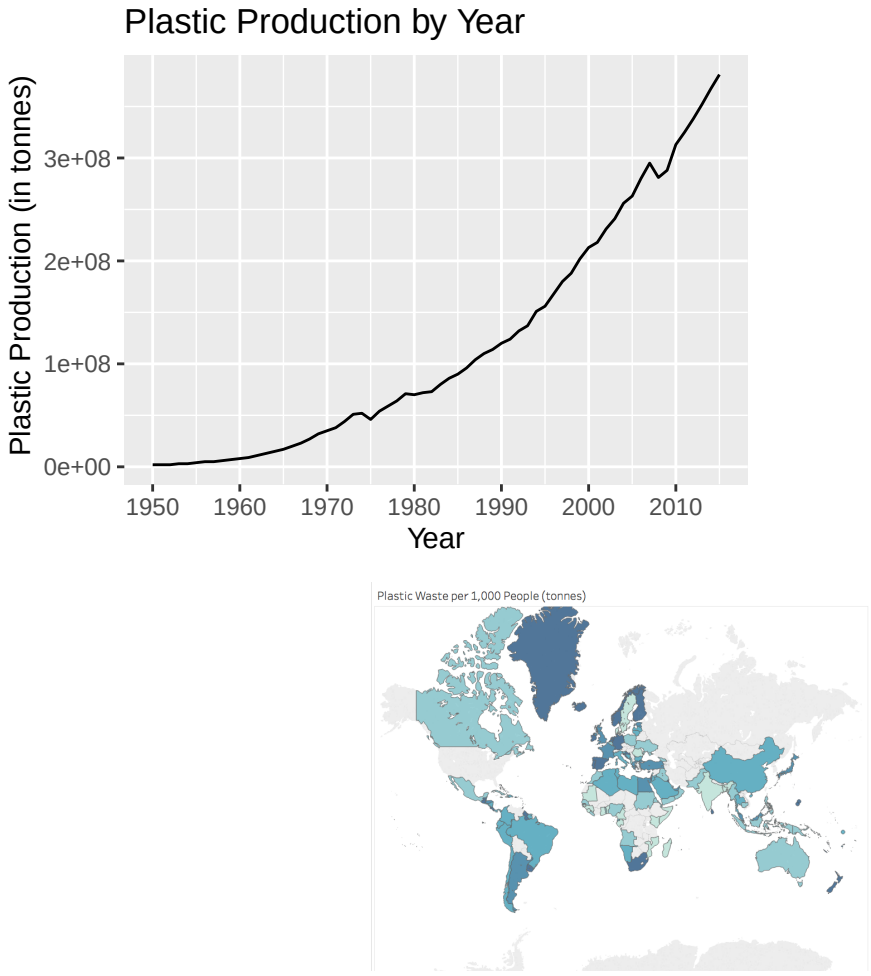


Figure 7: 2010 Worldmap by Plastic Production (population-standardized)

I used the Global Plastics Production over Time and the 2010 Plastic Production by Country datasets from Our World in Data in conjunction with 1950-2020 population breakdowns from the UN's data archive. I wanted to investigate the plastic waste:population ratio by country in 2010, as well as for the world at

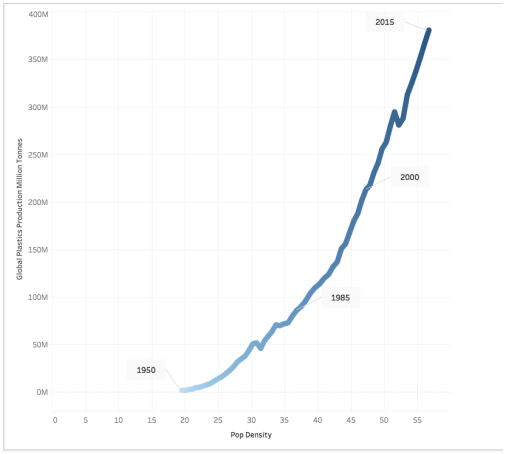


Figure 8: Time-Based Scatterplot of Population Density vs. Total Plastic Production.

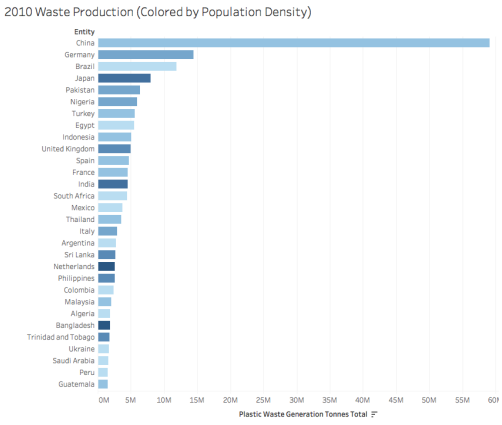


Figure 9: Ordered Bar Chart of Countries (with nonmissing data) colored by relative population density.

large from 1950 to the present I merged these datasets together and recreated a simple lineplot showing the exponential growth of world plastic production in the last few decades. I also regressed few basic population indicators on plastic production. **Interesting trends:**

- plastic waste production far outpaces population growth, particularly in recent years.
- though total population showed a significant association with total plastic production in 2010 (which makes a lot of sense) it was surprising that the countries with the densest populations weren't necessarily the ones with the highest total plastic production. For example: India has one of the world's biggest populations but was ranked lower than Germany, Spain, and Nigeria in Total Plastic Production in 2010. I'd like to use the UN population data (which breaks countries up by development status) to investigate this finding more. The final point I intend to convey with my part of the visualization is that plastic production has grown relatively independently of raw population growth, in the sense that it is not just the world's most populous countries contributing the most to global plastic production.

Micro and Macroplastic Pollution in the Ocean

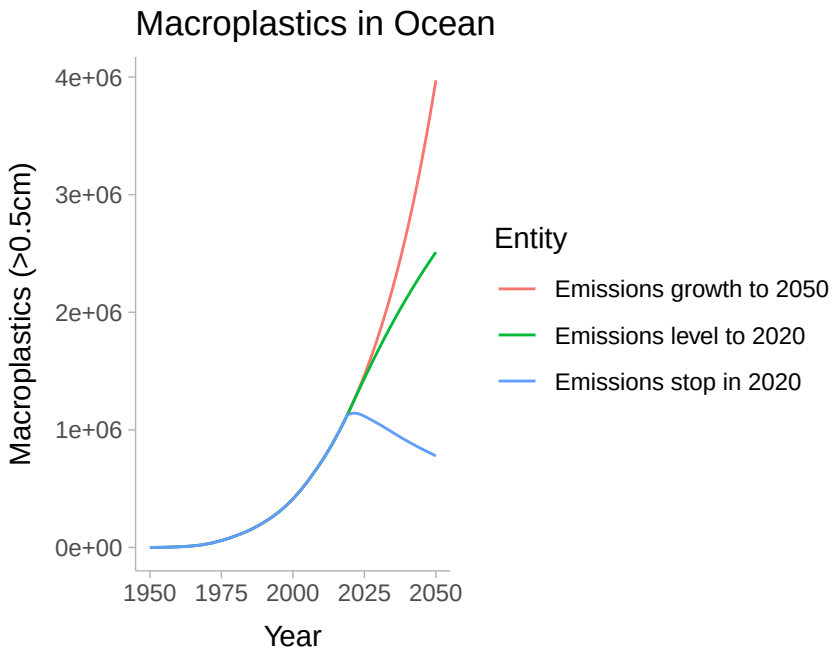
Both the Micro and Macroplastics Datasets are stored with the same organization. There are three categories of data: Plastic Pollution up until 2050 if emissions continue to increase at the projected rate, pollution until 2050 if the emission rate levels to the 2020 rate by 2050, and pollution until 2050 if the emissions rate levels off stop in 2020. Scroll below to see information about the data and our analyses.

Macroplastics

Show 10 entries Search:

	Entity	Missingness	Number of Observations	Year	Plastic Count Range
1	3 Types: Emissions Growth, Consistency, and Stop from 2020	No Missingness	303	1950 to 2050	0 to 3971900

Showing 1 to 1 of 1 entries Previous 1 Next



Microplastics

Show 10 entries Search:

	Entity	Missingness	Number of Observations	Year	Plastic Count Range
1	3 Types: Emissions Growth, Consistency, and Stop from 2020	No Missingness	303	1950 to 2050	0 to 2652700

Showing 1 to 1 of 1 entries Previous 1 Next

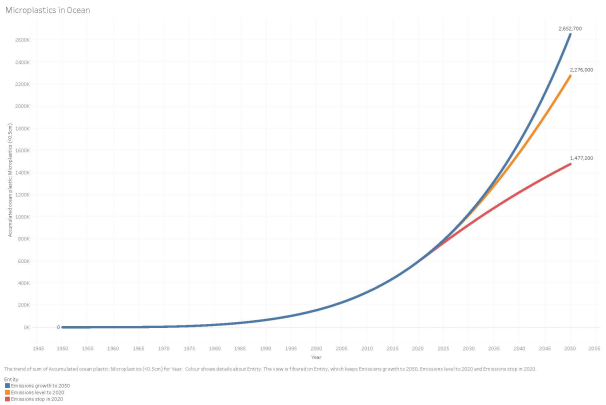


Figure 10: Microplastics

Macro and Microplastics Comparisons

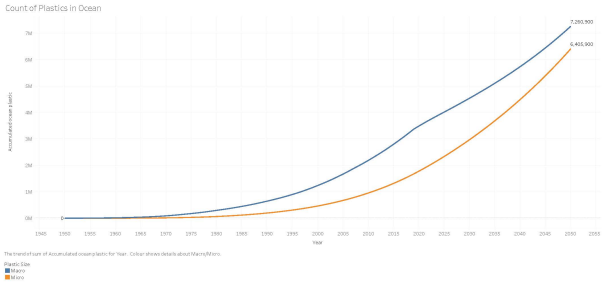


Figure 11: Micro vs. Macro

In comparing Micro and Macroplastics, we discovered that there are currently more Macroplastics polluting the ocean than Microplastics. Interestingly, if plastic emissions were to stop in 2020, the number of Macroplastics is projected to drop significantly to be less than the number of Microplastics by 2050.

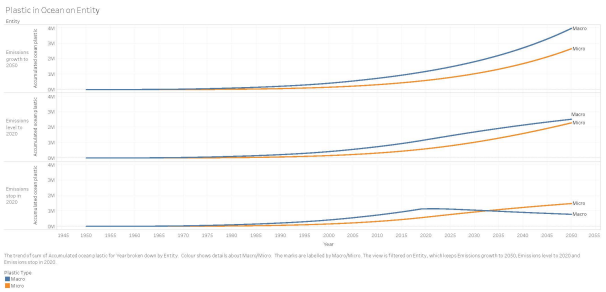


Figure 12: Micro and Macro

After generating the graphs above we reached several conclusions. Overall, for both Macro and Microplastics, the current rate of emissions is projected to lead to exponential growth in the amount of plastic pollution in the ocean. This growth is projected to curb that growth if emissions level to their 2020 levels by 2050, but the only real way to decrease the amount of ocean plastics is if emissions stop completely.

Conclusion

Notable features we have the option to visualize are as follows:

- Translations between polymer types to more commonly understood metrics for consumers. Polymer type waste of all kinds has increased over time.
- Plastic waste production far outpaces population growth, particularly in recent years.
- Countries with the densest populations weren't necessarily the ones with the highest total plastic production.
- Micro and Macroplastic levels in the ocean are growing at an exponential rate, and will continue to do so if pollution levels both remain at their current level and continue to grow as projected.

Link to Github Repo

Proof that we did in fact write the code to present this data :-) Although many of the graphs are simply images dropped in from Tableau.

https://github.com/kylierlin/Infovis_FP