Final Report: Scenario 2 - MPX in the EU

Dawson Coblin and Jose Lazo

2022-12-11

Problem Statement

As the world begins to recover from the effects of COVID, the threat of the monkeypox virus (MPX) has public health officials preparing for the worst. Throughout Europe (EU), cases of MPX have begun to rise, and it is imperative to understand the groups at the highest risk for infection. This report aims to prepare data to analyze and assess the impact of MPX within each country of the EU, understand how case rates may differ by region, and find potential links to various demographic factors. These findings will aid local health departments' implementation and appropriate response to the epidemic. Our team hypothesizes that MPX will be prevalent among young (under 15 years old) and older populations (over 64 years old), as is typical for many infectious diseases. We believe that countries with larger proportions of young and old populations will have more cases of MPX than those regions that have countries with higher middle aged population proportions.

Methods

To prepare this report our team uses data sets for the EU, explicitly including information on confirmed cases of MPX per country (euro_mpx_cases.csv) and census data (euro_census_stats.csv). The trajectory of this investigation incorporated four sets of data manipulation: (1) Grouping countries into regions; (2) Assessing necessary census information; (3) Compiling data sets with confirmed case demographics; (4) Creating tables and figures for visual analysis.

Countries were grouped by region based on a region's data set (world_country_regions.csv) to enhance the representation of positive cases per region. Within the euro_mpx_cases data set, relevant fields include country, country code, and confirmed cases. A curated demographic data set included information on age, sex, and the number of people per strata. For the ecologic analysis of monkeypox and age, countries with missing census data for individual age stratum were omitted from analysis.
As an aside with the census data, we chose to initially assess any "missing" data as "NA" when paired with a country. However then decided to print information with omit-NA's to allow aggregate fields to be added together for review of population demographic (Population Density) per Region"

Results

Table 1: MPX Cases Per Region

EU Region	Total Cases Per Region	Total Region Population	Countries Per Region	Case Rate per 100k
Eastern Europe	277	89171711	6	0.31
Northern Europe	619	37518701	9	1.65
Southern Europe	7958	133879004	7	5.94
Western Asia	4	904705	1	0.44
Western Europe	8214	191156200	6	4.29

In Table 1, we see that the regions in EU include various numbers of countries, thus having unequal populations. A useful metric to create a graph will be the calculation of cases per 100,000. This will allow us to use one categorical variable (EU Region) to a quantitative variable (number of confirmed cases). We compiled the relevant confirmed case information per region in the EU with the cumulative case rate per 100k.

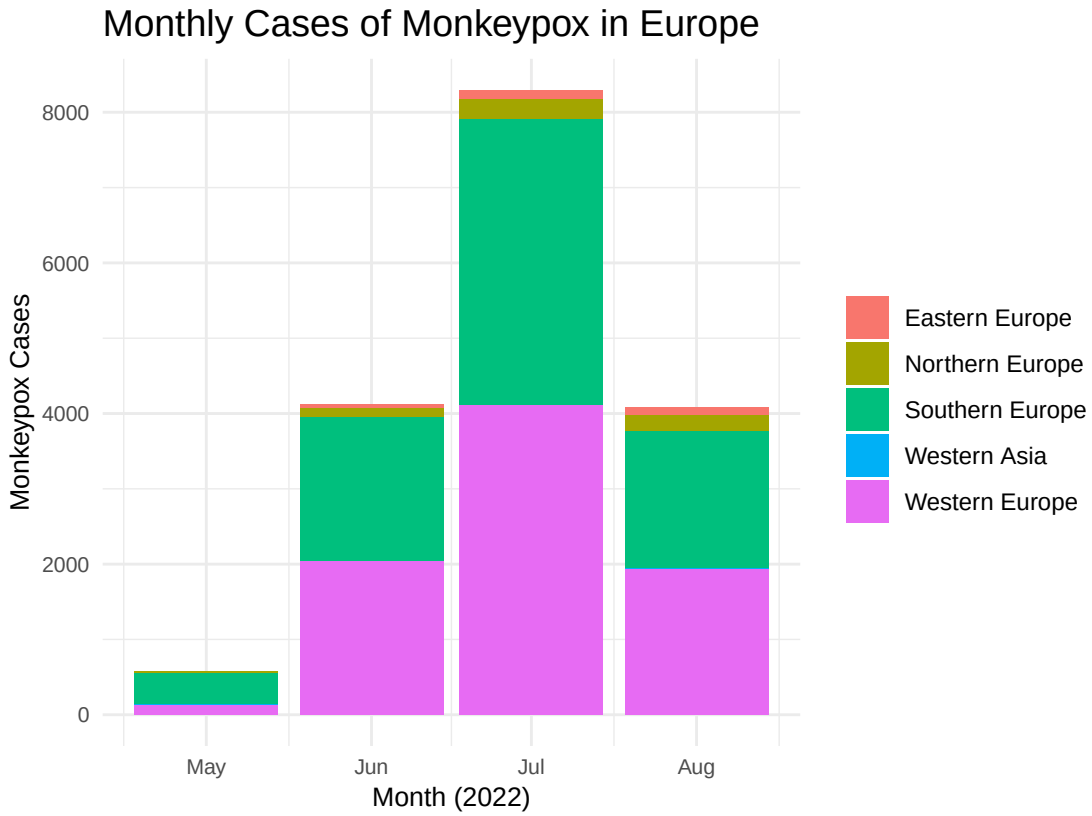

Figure 1

In the bar graph (Fig. 1) above we see the breakdown of monthly MPX infections in Europe between May and August of this year. The majority of cases occurred in Western Europe and Southern Europe as evidenced by the respectively colored purple and green regions in the bar graph. Over the course of 4 months, the total number of MPX cases by region were Western Asia = 4, Eastern Europe = 277, Northern Europe = 619, Southern Europe = 7958, and Western Europe = 8214

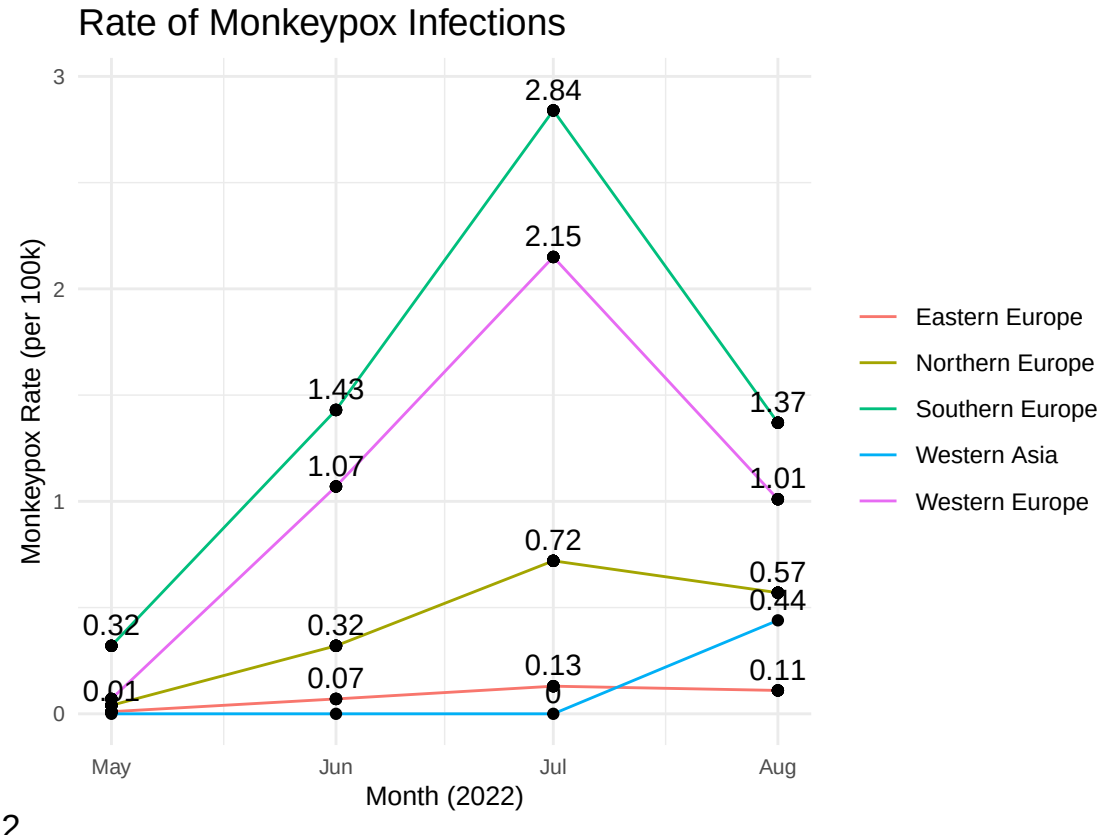

Figure 2

As a follow up to the bar graph (Fig. 1) above, we wanted to look at the rate of MPX infections per region. In the bar graph above, Western Europe had the greatest number of MPX infections however when we look at the rate of infections (Fig. 2), we see that over the course of the four months, Southern Europe had the greatest rate of infection with a peak of 2.84 MPX cases per 100,000 people in the month of July.

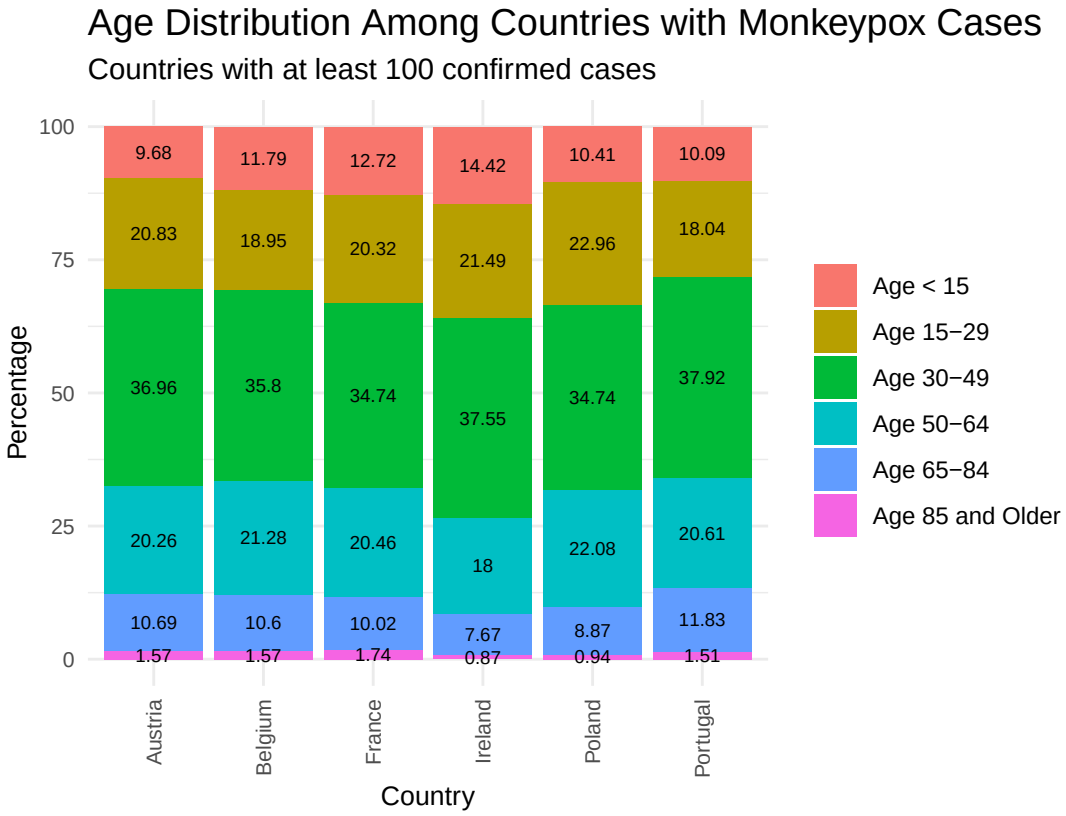

Figure 3A

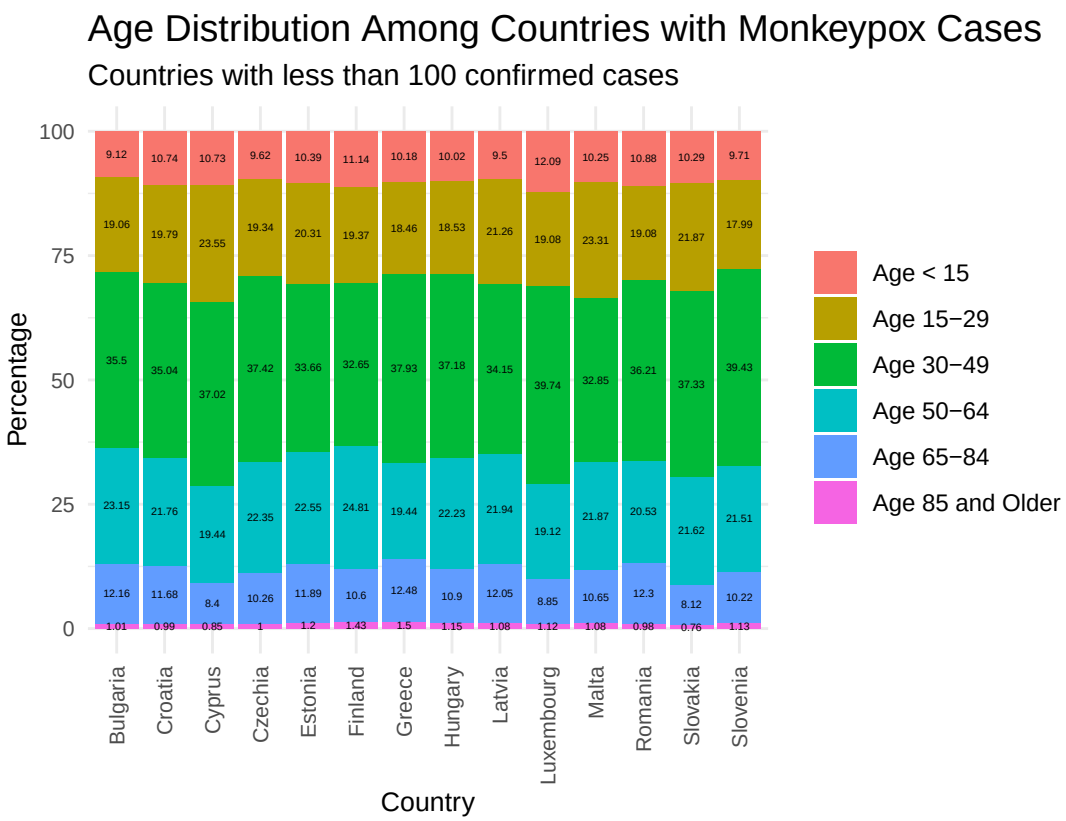

Figure 3B

In our dataset, 9 countries were missing population data for age stratums which left us with 20 countries with MPX cases ranging 4 - 2889 and a median of 42. Figure 3 includes two plots dividing countries that experienced at least 100 cases of MPX with those that had less than 100 cases. The figure parameters include stacked bar graphs that visually represent the age brackets from which a positive case is observed. When comparing the two graphs, the proportion among each age stratum appear similar with the stratum of age < 15 ranging from 9.68%-14.42% and 9.12% - 12.09% in countries with $>= 100$ cases and countries with $<$ 100 cases, respectively. Similar proportions were also seen for age stratum > 64. The 30-49 age range makes up the largest proportion of population among the 20 European country in which a MPX case was reported and age stratum information was available. In a seperate analyses of countries with confirmed MPX and sex (M/F), no relationship could be seen with country level sex breakdown and number of MPX cases.

Discussion

Our analysis of Table 1 and Figure 1 reveal the impacted regions for MPX though the study period (May through August 2022) occurred in the Southern and Western Regions of Europe. Within the EU, these two regions comprise 13 of its 29 total countries, roughly 71.8% of the total EU population (extrapolated using Table 1).

Through analysis of Figure 3, we are not confident of the initial hypothesis regarding younger (under 15 years old) and older (over 64 years old) populations proportions of country demographics causing a higher risk of infection. Review of the demographics associated to Figure 3 led to an inconclusive relationship between age and MPX.

Given the analysis of this information, health department leadership should begin educational programs and outreach focusing on countries in Western and Southern Europe.