

June 2021

# FOOD and COVID-19 RELATIONSHIPS

Data Visualization Project

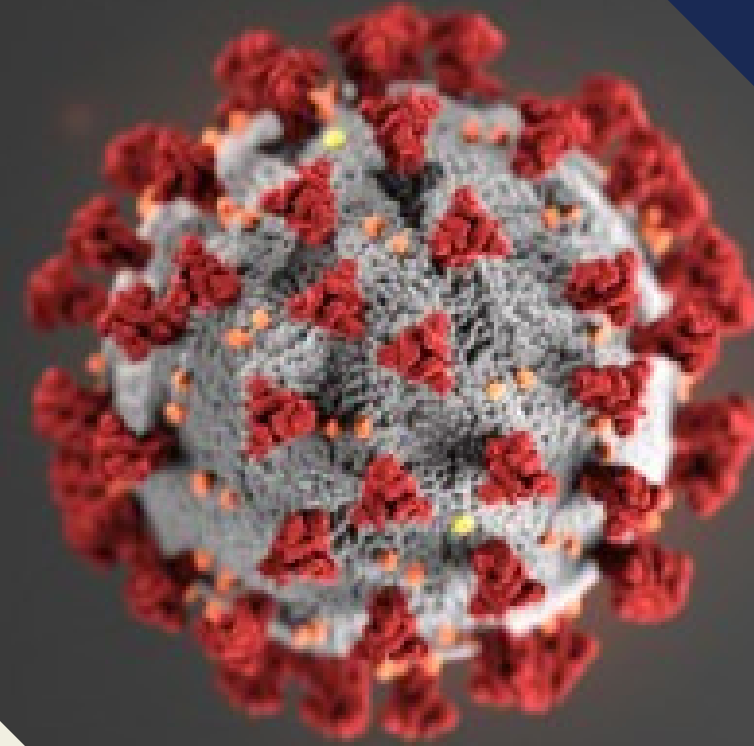
Presented By  
**Furkan Mert Şahin**  
**16030411042**

# Introduction

Due to the coronavirus epidemic that swept the whole world in 2020, millions of people caught this disease and millions of people lost their lives unfortunately.

There are many factors that affect people who die from coronavirus disease. Factors such as eating habits, smoking habits, age, and chronic disease affected the death and recovery rates in this disease.

Undoubtedly, nutritional habits and valuable nutrients from food are very important for human health, and this situation has become more important during the coronavirus process.





# COVID-19 & Food Datasets

The data set shows the percentage of products consumed by countries and continents. The dataset includes food groups such as alcoholic products, animal products, cereals, sugar and vegetables. The total number of data in the data set is 144.

```
# ----- Food Supply Quantity (kg) Intake of Countries Dataset -----  
  
## This dataset includes percentage of food intake (kg) in countries around the world.  
  
country_food_supply = pd.read_csv("Food_Supply_Quantity_kg_Data.csv");  
country_food_supply = pd.merge(country_food_supply,worldmeter_coronavirus, on = "Country")  
country_food_supply = country_food_supply.drop(["Unit (all except Population)", "Miscellaneous", "Undernourished", "Offals", "Starchy R",  
"Obesity", "Treenuts", "Oilcrops", "Meat", "Stimulants", "Sugar Crops", "Fruits - Excludin",  
"Vegetal Products", "Vegetable Oils", "Eggs", "Confirmed", "Deaths", "Active"], ax  
country_food_supply
```

	Country	Alcoholic Beverages	Animal Products	Cereals - Excluding Beer	Sugar & Sweeteners	Vegetables	Recovered	Population	Continent
0	Afghanistan	0.0014	9.4341	24.8097	1.3489	6.7642	0.123374	38928000.0	Asia
1	Albania	1.6719	18.7684	5.7817	1.5367	11.7753	1.792636	2838000.0	Europe
2	Algeria	0.2711	9.6334	13.6816	1.8342	11.6484	0.167572	44357000.0	Africa
3	Angola	5.8087	4.9278	9.1085	1.8495	2.3041	0.056808	32522000.0	Africa
4	Argentina	4.2672	19.3454	8.4102	3.0536	4.3503	3.905192	45377000.0	South America
...	...	...	...	...	...	...	...	...	...
140	Vanuatu	0.2424	6.6244	6.4102	1.3645	3.9175	0.000312	321000.0	Australia/Oceania
141	Vietnam	1.4591	8.5765	16.8740	1.2846	11.9508	0.001526	96209000.0	Asia
142	Yemen	0.0364	5.7874	27.2077	5.0468	3.2135	0.004788	29826000.0	Asia
143	Zambia	5.7360	6.0197	21.1938	1.5632	3.4649	0.290524	18384000.0	Africa
144	Zimbabwe	4.0552	8.1489	22.6240	4.6485	2.3213	0.190964	14863000.0	Africa

# COVID-19 & Food Datasets

Another dataset of the data visualization project shows how much protein countries and continents get from the food groups they consume. The dataset includes food groups such as alcoholic products, animal products, cereals, sugar and vegetables. The total number of data in the data set is 144.

```
# ----- Protein Intake of Countries Dataset -----  
  
## This dataset includes percentage of protein intake from different types of food in countries around the world.  
  
country_protein_supply = pd.read_csv("Protein_Supply_Quantity_Data.csv");  
country_protein_supply = pd.merge(country_protein_supply,worldmeter_coronavirus, on = "Country")  
country_protein_supply = country_protein_supply.drop(["Unit (all except Population)","Miscellaneous","Undernourished","Offals","Sta  
"Obesity","Treenuts","Oilcrops","Meat","Stimulants","Sugar Crops","Fruits - Excludin  
"Vegetal Products","Vegetable Oils","Eggs","Confirmed","Deaths","Active"], ax  
  
country_protein_supply
```

	Country	Alcoholic Beverages	Animal Products	Cereals - Excluding Beer	Sugar & Sweeteners	Vegetables	Recovered	Population	Continent
0	Afghanistan	0.0000	9.7523	35.9771	0.0000	1.1370	0.123374	38928000.0	Asia
1	Albania	0.1840	27.7469	14.2331	0.0042	3.2456	1.792636	2838000.0	Europe
2	Algeria	0.0323	13.8360	26.5633	0.0000	3.1267	0.167572	44357000.0	Africa
3	Angola	0.6285	15.2311	20.3882	0.0092	0.8133	0.056808	32522000.0	Africa
4	Argentina	0.1704	31.9799	13.6702	0.0049	1.0516	3.905192	45377000.0	South America
...	...	...	...	...	...	...	...	...	...
140	Vanuatu	0.0078	19.8713	14.6387	0.0000	1.5847	0.000312	321000.0	Australia/Oceania
141	Vietnam	0.1555	20.4466	18.5247	0.0056	3.7216	0.001526	96209000.0	Asia
142	Yemen	0.0000	10.0122	35.1179	0.0000	0.5448	0.004788	29826000.0	Asia
143	Zambia	0.4824	9.8925	28.5182	0.0000	0.8039	0.290524	18384000.0	Africa
144	Zimbabwe	0.2929	11.3443	33.1934	0.0000	0.5955	0.190964	14863000.0	Africa

# COVID-19 & Food Datasets

The final dataset of the data visualization project shows how much energy (kcal) countries and continents get from the food groups they consume. The dataset includes food groups such as alcoholic products, animal products, cereals, sugar and vegetables. The total number of data in the data set is 144.

```
# ----- Energy Intake of Countries (kcal) Dataset -----  
  
# This dataset includes percentage of energy intake (kcal) from different types of food in countries around the world.  
  
country_kcal_supply = pd.read_csv("Food_Supply_kcal_Data.csv");  
country_kcal_supply = pd.merge(country_kcal_supply,worldmeter_coronavirus, on = "Country")  
country_kcal_supply = country_kcal_supply.drop(["Unit (all except Population)", "Miscellaneous", "Undernourished", "Offals", "Starchy Re",  
                                                "Obesity", "Treenuts", "Oilcrops", "Meat", "Stimulants", "Sugar Crops", "Fruits - Excluding",  
                                                "Vegetal Products", "Vegetable Oils", "Eggs", "Confirmed", "Deaths", "Active"], axis=1)  
country_kcal_supply
```

	Country	Alcoholic Beverages	Animal Products	Cereals - Excluding Beer	Sugar & Sweeteners	Vegetables	Recovered	Population	Continent
0	Afghanistan	0.0000	4.7774	37.1186	2.2261	0.7504	0.123374	38928000.0	Asia
1	Albania	0.9120	16.0930	16.2107	3.4422	2.7508	1.792636	2838000.0	Europe
2	Algeria	0.0896	6.0326	25.0112	3.9869	2.0457	0.167572	44357000.0	Africa
3	Angola	1.9388	4.6927	18.3521	2.7539	0.3525	0.056808	32522000.0	Africa
4	Argentina	1.4354	14.9869	16.7927	7.0536	0.8643	3.905192	45377000.0	South America
...	...	...	...	...	...	...	...	...	...
140	Vanuatu	0.2668	7.4519	16.1044	2.4204	0.6099	0.000312	321000.0	Australia/Oceania
141	Vietnam	0.7150	10.9806	26.9833	2.0940	1.9578	0.001526	96209000.0	Asia
142	Yemen	0.0000	3.4667	32.0727	7.0303	0.3636	0.004788	29826000.0	Asia
143	Zambia	1.1925	3.3043	31.5528	2.5590	0.4472	0.290524	18384000.0	Africa
144	Zimbabwe	1.4269	3.9356	29.8044	5.0403	0.2532	0.190964	14863000.0	Africa

# Preparation of COVID-19 & Food Datasets

```
country_food_supply.isnull().sum()    # Columns with Missing Data in the Food Supply Dataset

Country                               0
Alcoholic Beverages                   0
Animal Products                       0
Cereals - Excluding Beer              0
Sugar & Sweeteners                    0
Vegetables                           0
Recovered                             3
Population                            0
Continent                             0
dtype: int64
```

*Columns with Missing Data in the Food Supply Dataset*

```
country_protein_supply.isnull().sum()  # Columns with Missing Data in the Protein Supply Dataset

Country                               0
Alcoholic Beverages                   0
Animal Products                       0
Cereals - Excluding Beer              0
Sugar & Sweeteners                    0
Vegetables                           0
Recovered                             3
Population                            0
Continent                             0
dtype: int64
```

*Columns with Missing Data in the Protein Supply Dataset*

COVID-19 & Food datasets need to go through a data preparation process prior to visualization. Missing observation analysis was used in the preliminary preparation of the data. There are 3 missing data in the column with the recoveries in the data set. There is no missing data in other columns.

# Preparation of COVID-19 & Food Datasets

```
# Food Supply Dataset

# Since the data gap in the columns was not too large, the missing data was filled with the value assignment approach.

country_food_supply["Recovered"].fillna(country_food_supply["Recovered"].mean(),inplace=True)

country_food_supply.isnull().sum() # With the value assignment approach, missing data were filled with average data.
```

```
Country          0
Alcoholic Beverages  0
Animal Products   0
Cereals - Excluding Beer  0
Sugar & Sweeteners  0
Vegetables        0
Recovered         0
Population        0
Continent         0
dtype: int64
```

*Filling in Missing Data in the Food Supply Dataset*

```
# Energy (kcal) Supply Dataset

# Note: Since there are the same columns in each dataset, the same operations were performed.

country_kcal_supply["Recovered"].fillna(country_kcal_supply["Recovered"].mean(),inplace=True)

country_kcal_supply.isnull().sum() # With the value assignment approach, missing data were filled with average data.
```

```
Country          0
Alcoholic Beverages  0
Animal Products   0
Cereals - Excluding Beer  0
Sugar & Sweeteners  0
Vegetables        0
Recovered         0
Population        0
Continent         0
dtype: int64
```

*Filling in Missing Data in the Energy (kcal) Supply Dataset*

In order to clean up the missing data in the columns and not to spoil the significance of the data set during this process, missing data were filled in by averaging the recovered column.

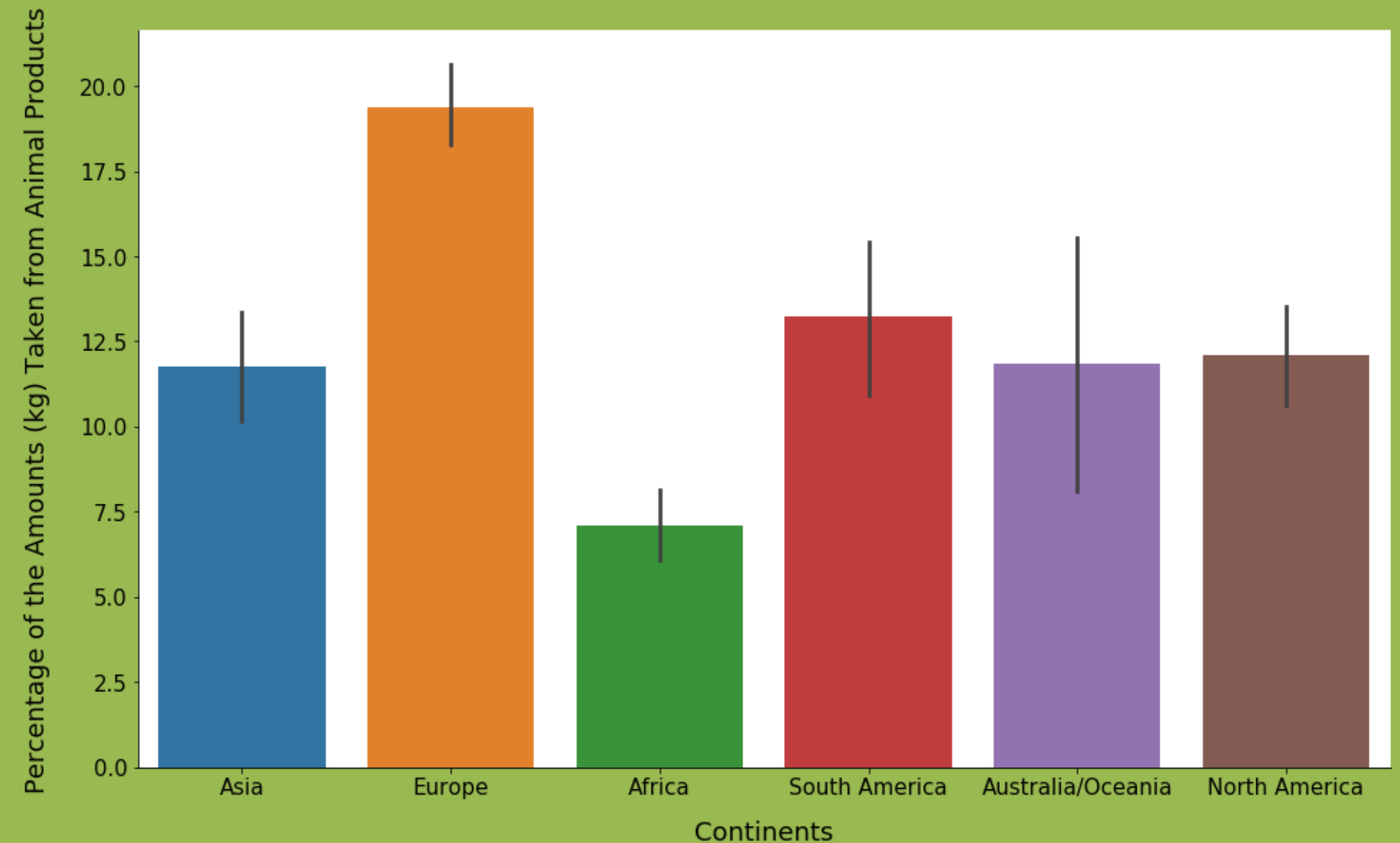
Since the number of missing data is low, filling with the mean approach did not impair the significance of the data sets. The filling process is done with the "fillna" method of the Pandas library, which is one of the Python libraries.



# Visualization of COVID-19 and Food Datasets

## 1) Animal Products

The bar chart shows the percentage of animal products consumed by the continents. In the graph, it is understood that the continent that consumes animal products the most is Europe and that the continent that consumes the least is Africa. Factors such as the amount of production and the financial situation of the society affect this picture.

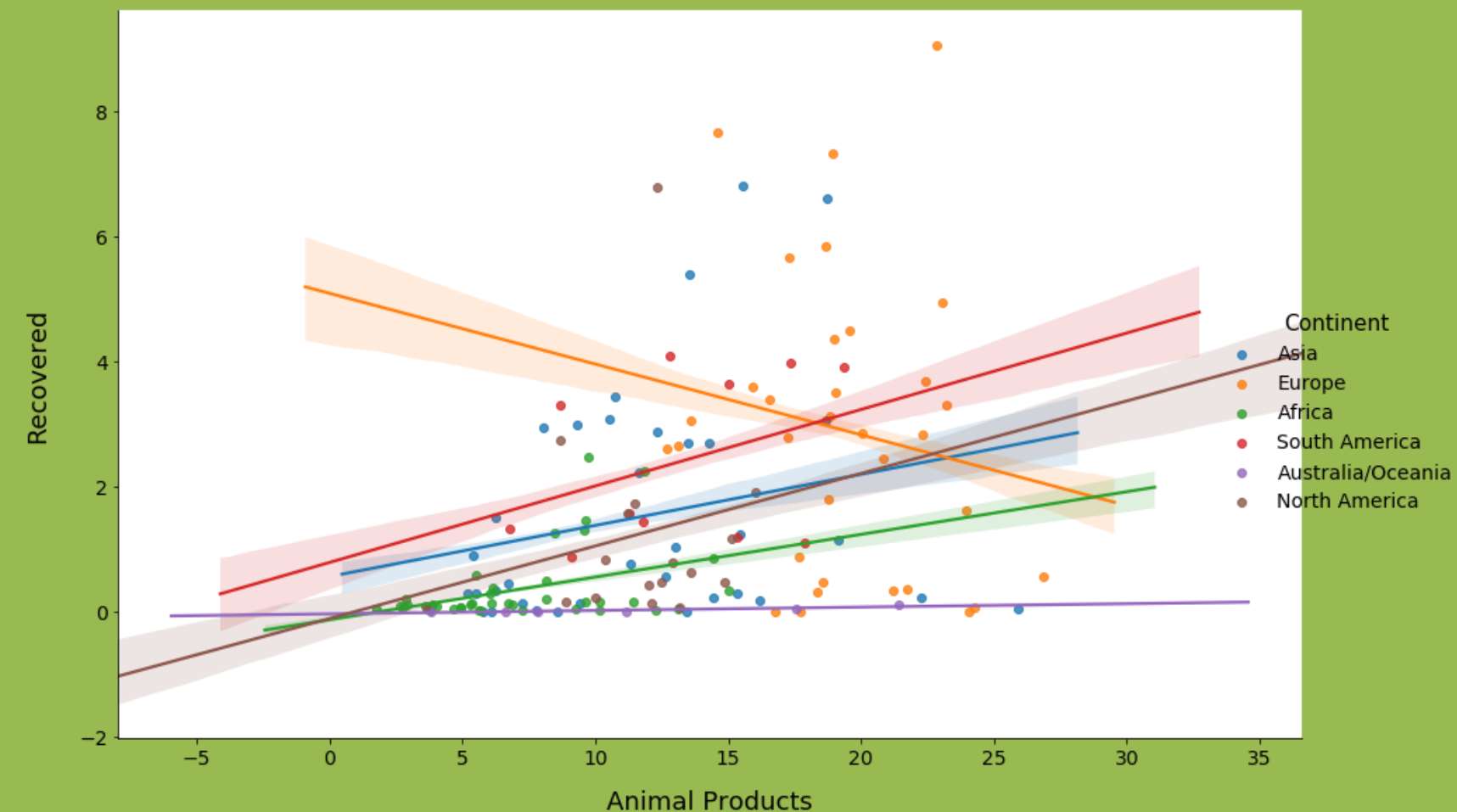
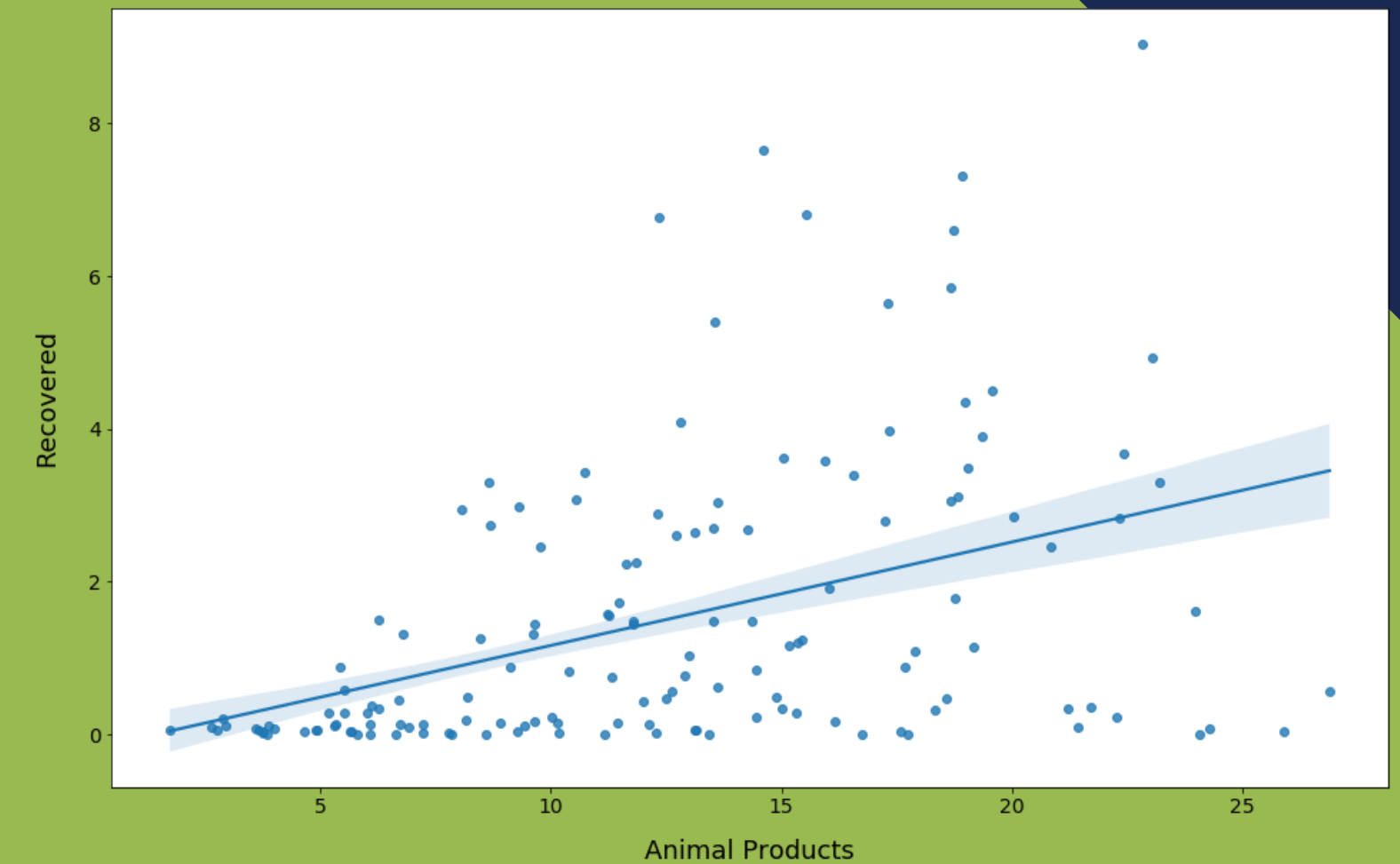




# Visualization of COVID-19 and Food Datasets

When the regression graph in the upper right is examined it is seen that the increase in animal product consumption has a positive contribution to recovery.

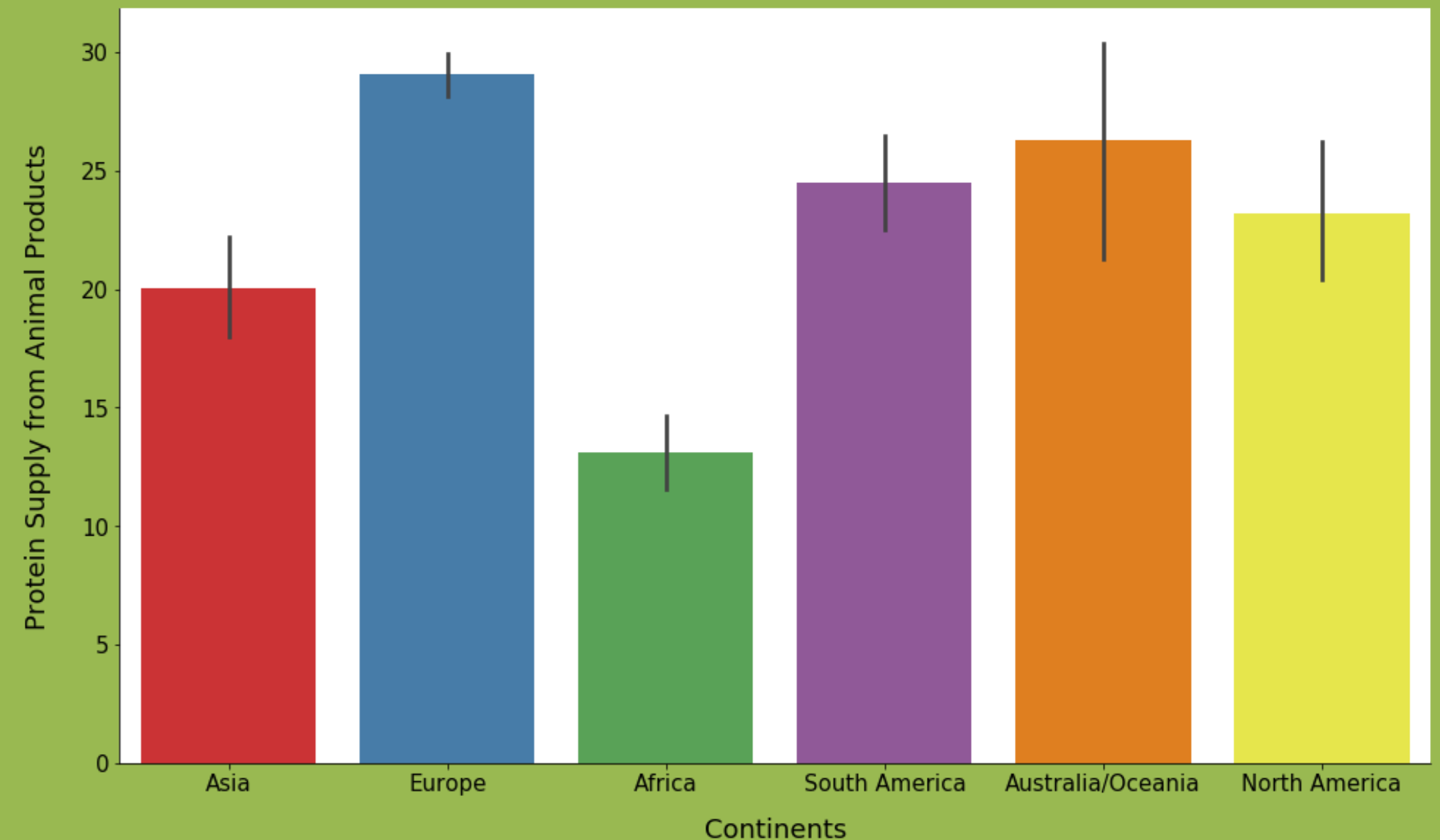
The other graph is a regression graph showing the effect of the amount of animal products consumed by the continents on recovery. The graph says the impact on recovery is negative as consumption of animal products increases in Europe.



# Visualization of COVID-19 and Food Datasets

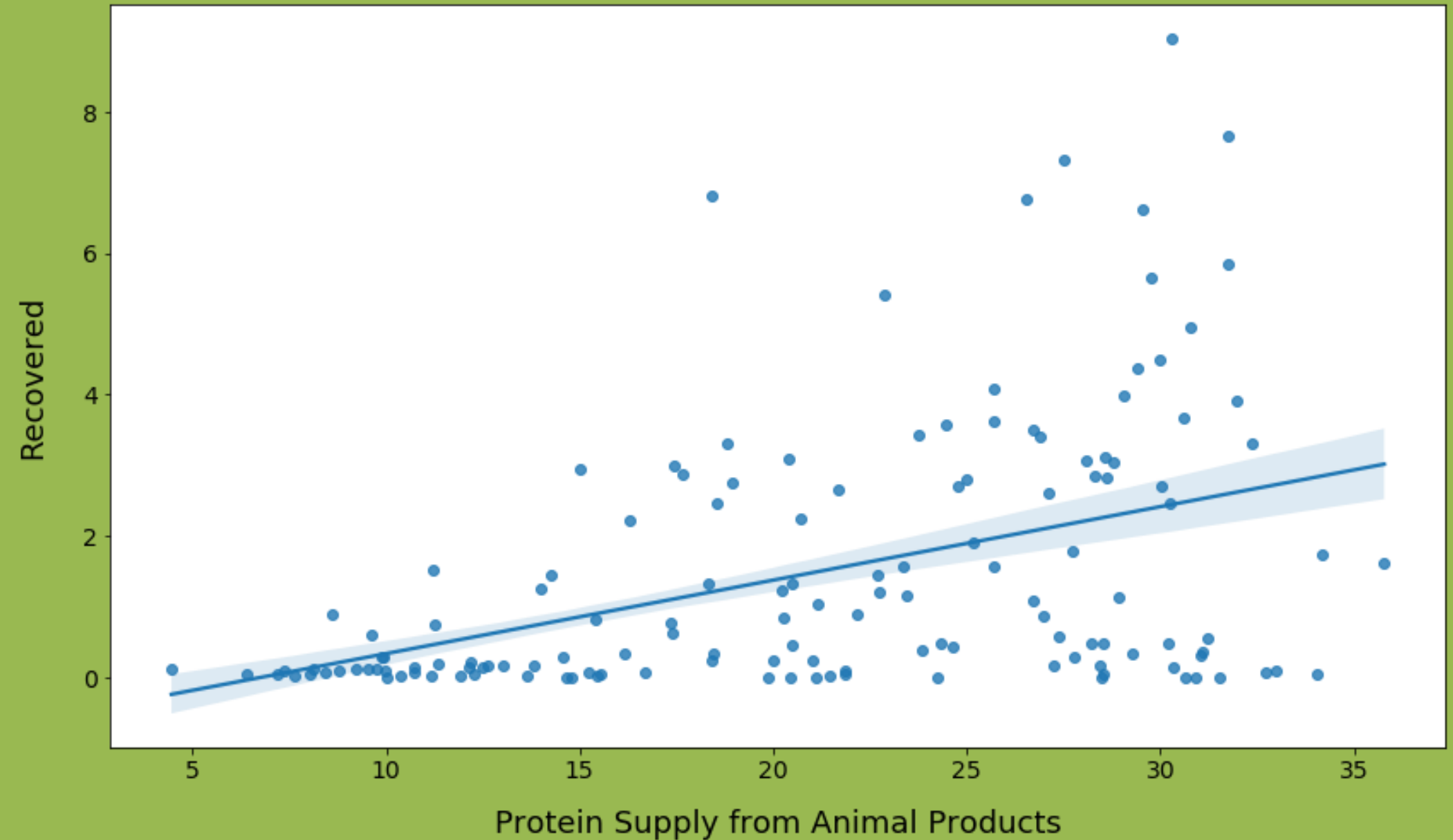
## 1a) Percentage of Protein Supply from Animal Products

Europe, which is the continent that consumes the most animal products, is also seen as the continent that consumes the most protein from animal products.



# Visualization of COVID-19 and Food Datasets

As the amount of protein supplied from animal products increases, the effect of healing increases in a positive sense. However, it cannot be said that the protein supply from animal products is the only factor in recovery. It can be observed that there is only a positive relationship.

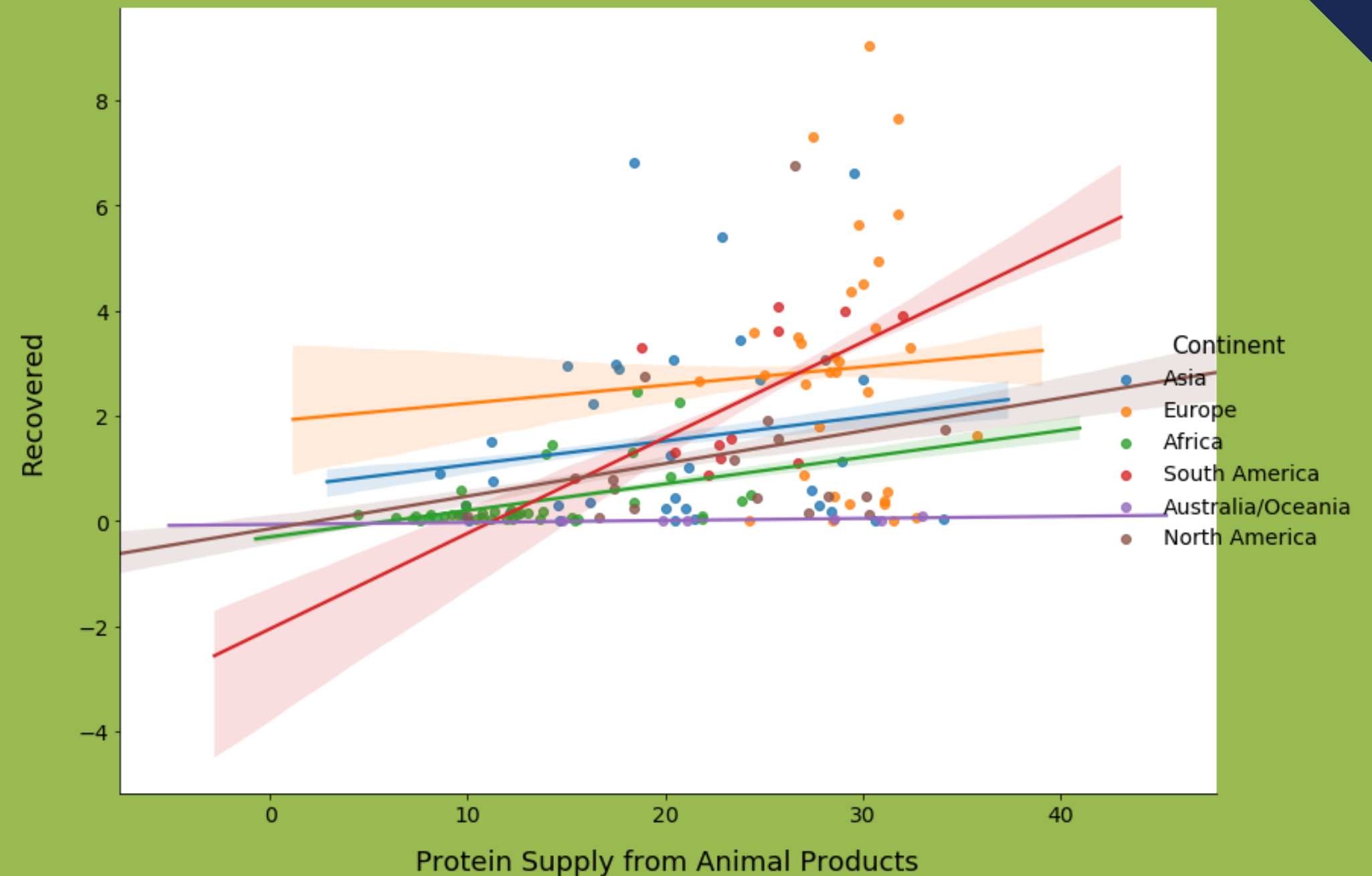


*The Relationship Between Recovered & Protein Supply from Animal Products*



# Visualization of COVID-19 and Food Datasets

When the effect of the protein supply from animal products on healing is examined on the basis of continents, the protein supply from animal products has increased the rate of people recovering in Europe, Africa, North America and Asia.



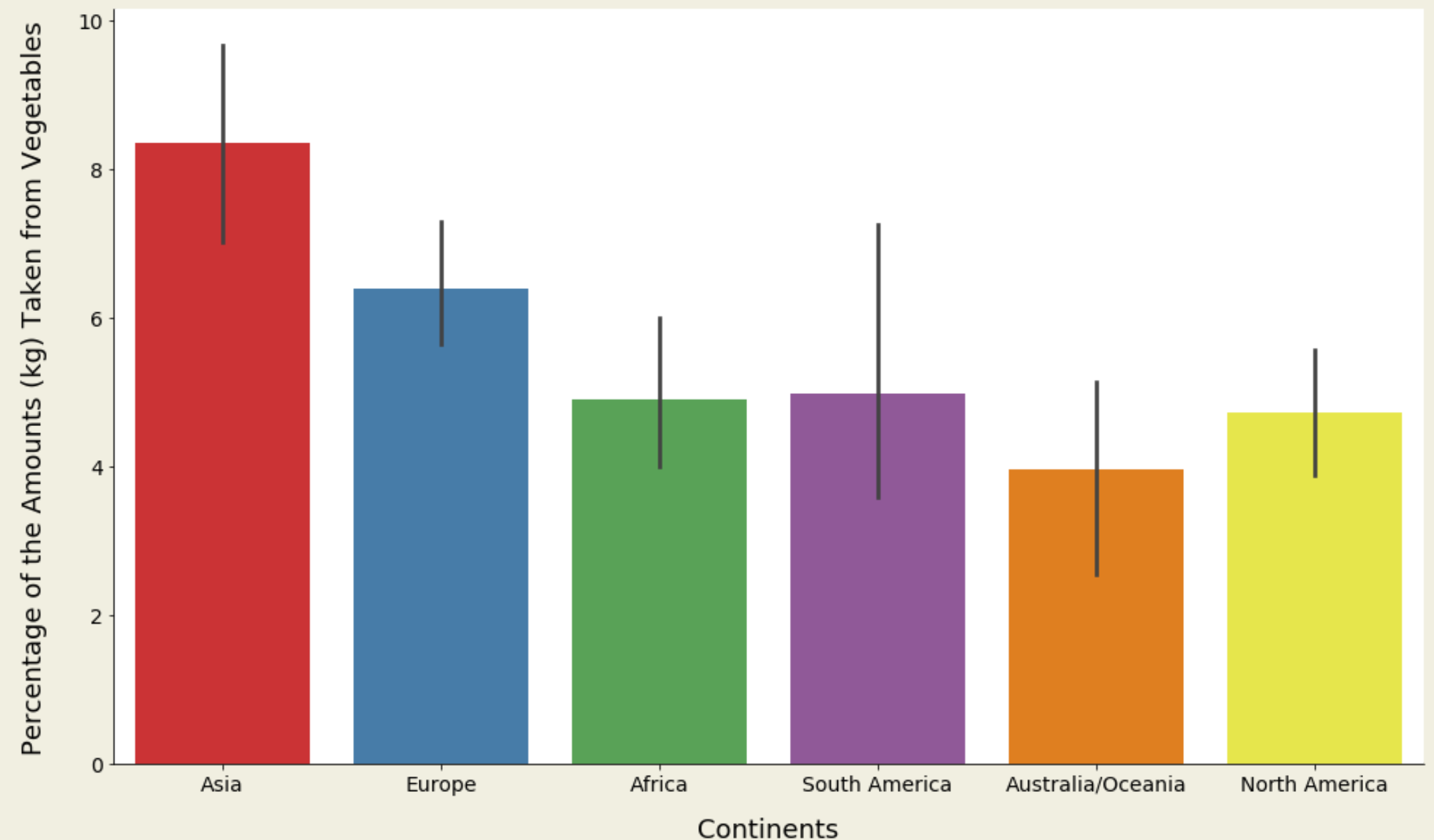
*The Relationship Between Recovered & Protein Supply from Animal Products by Continents*

# Visualization of COVID-19 and Food Datasets

## 2) Vegetables

When the bar graph of the vegetable category, which is one of the most important food groups in terms of human health, is examined, the continent that consumes the most vegetables is Asia.

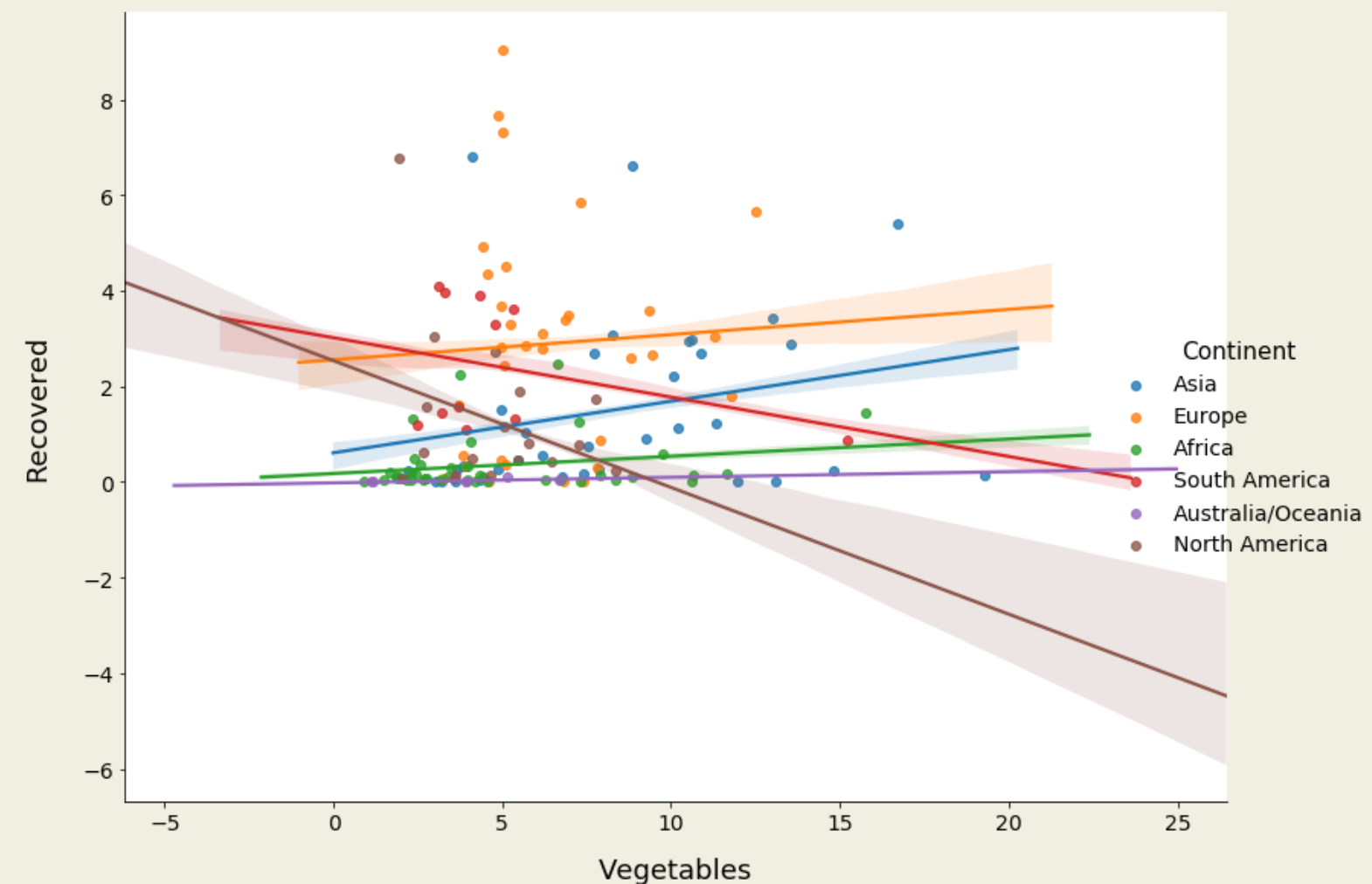
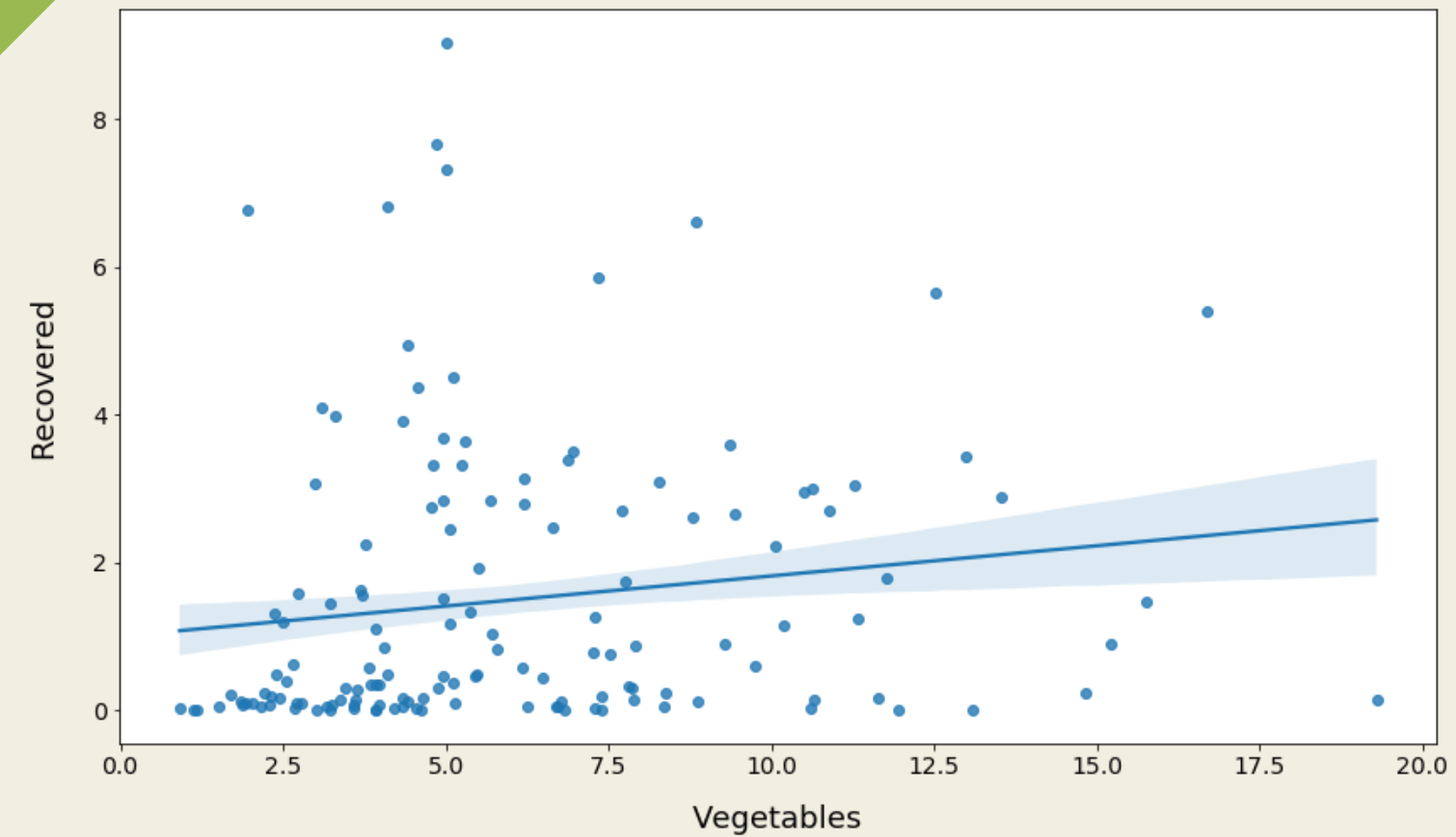
The continent that consumes the least vegetables is Australia/Oceania. Factors such as nutritional culture, fertile soil, agricultural production policy and purchasing power of the society may cause this graph to emerge.



# Visualization of COVID-19 and Food Datasets

In the regression graph, it is clear that there is a positive effect between vegetable consumption and recovery, but this relationship is not very strong because the slope of the regression line is low.

When examining the relationship between vegetable consumption and recovered on a continent basis, it is understood that recovery rates decrease when the amount of vegetable intake increases in North America and South America.

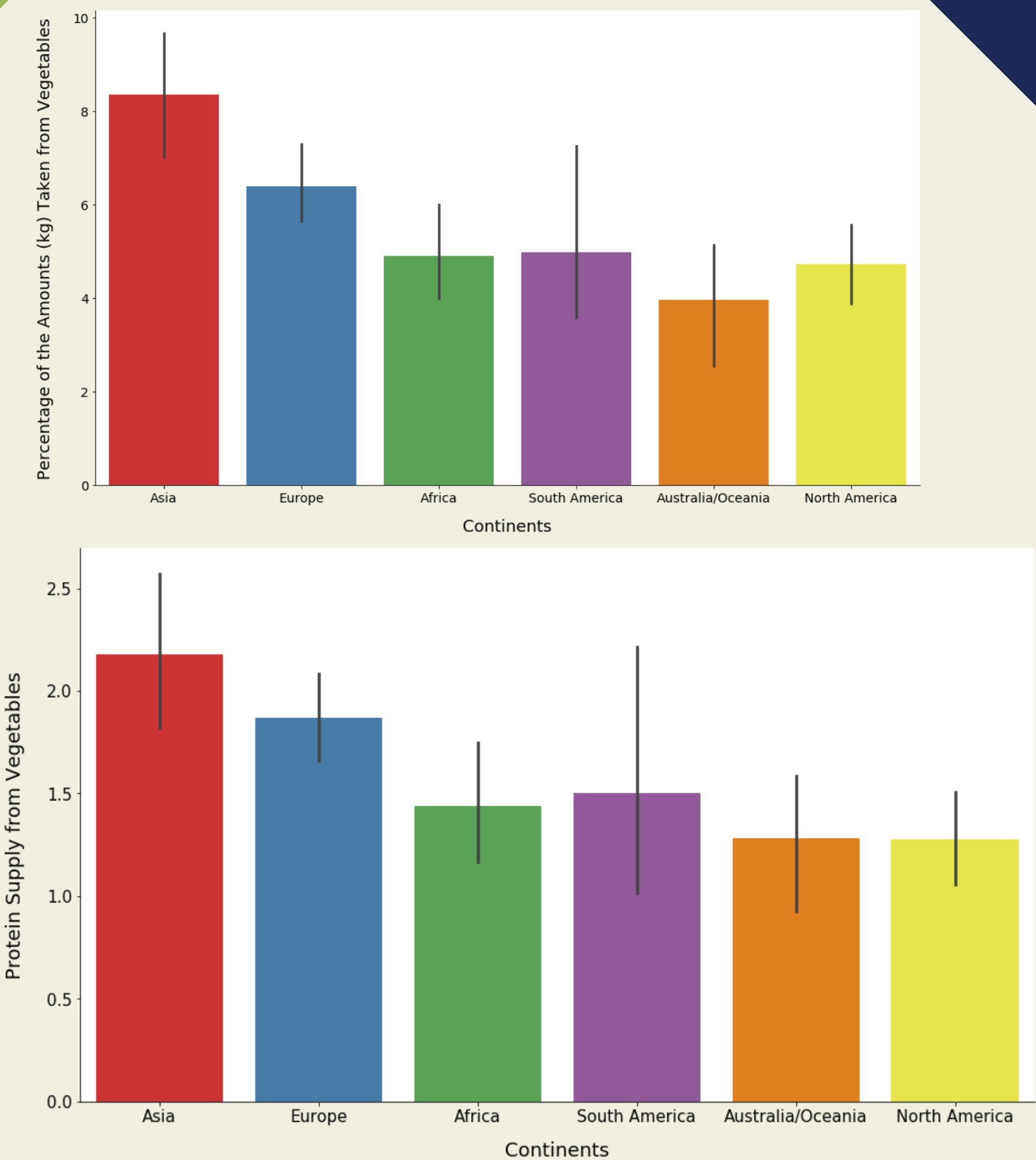




# Visualization of COVID-19 and Food Datasets

## 2a) Percentage of Protein Supply from Vegetables

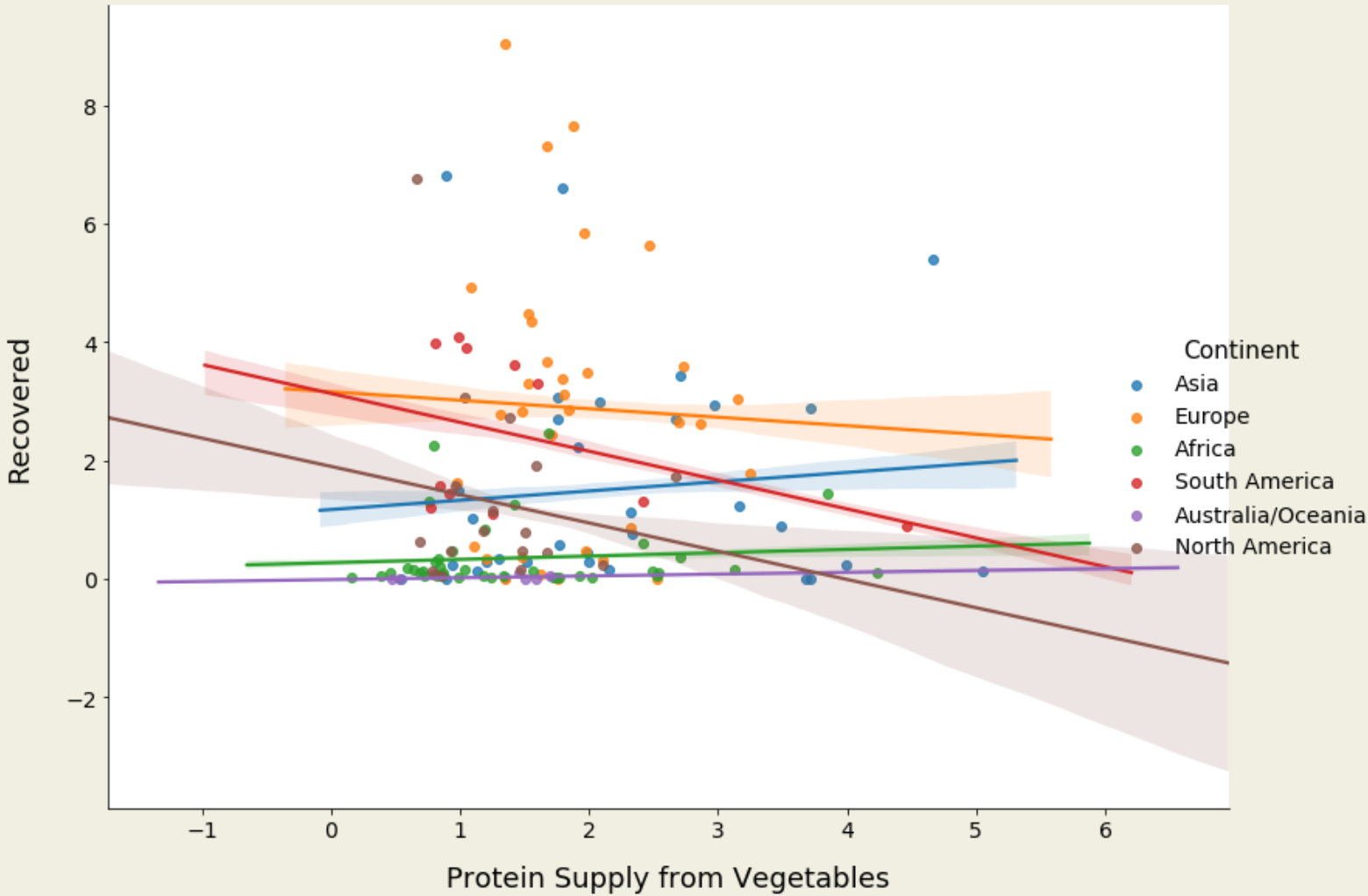
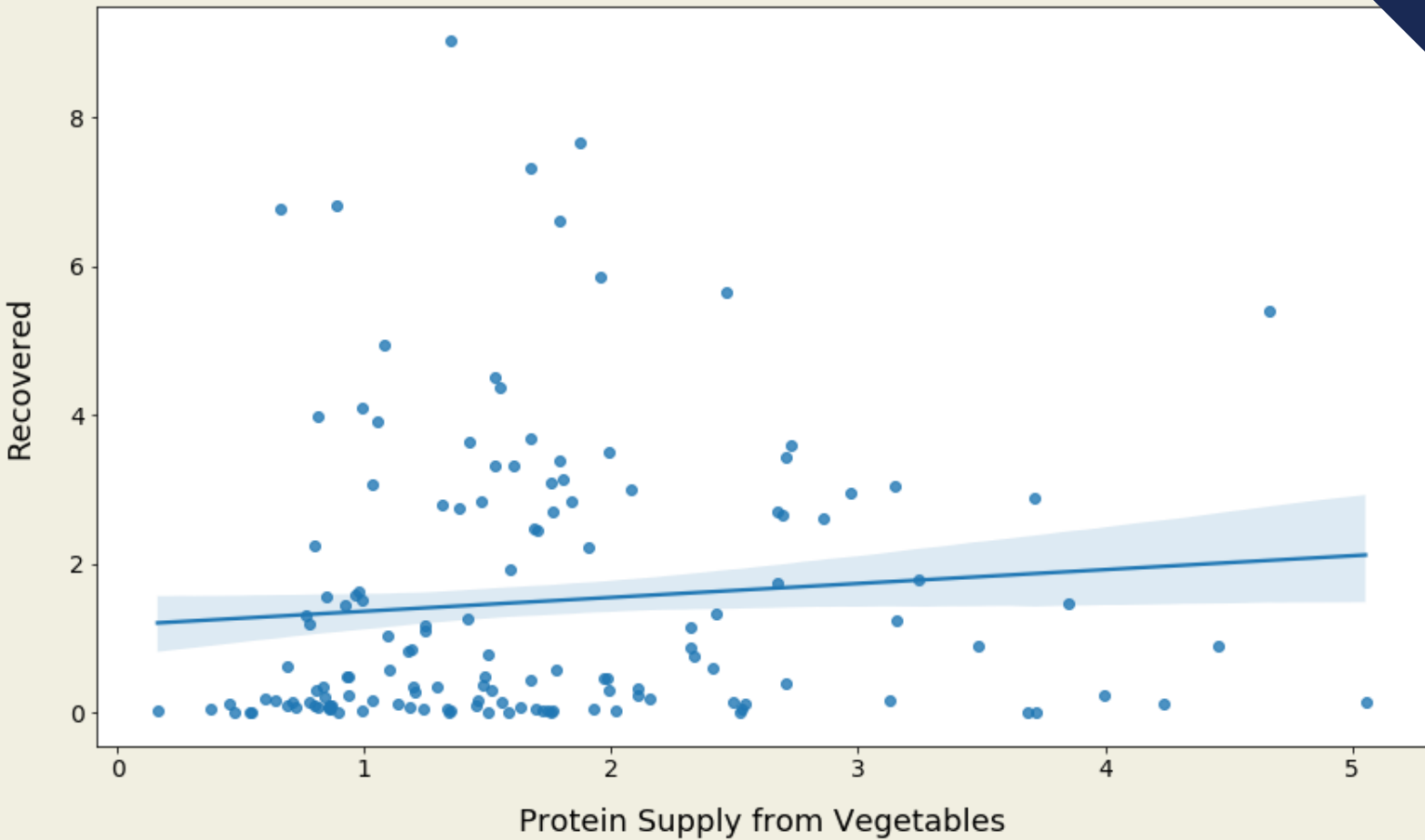
There is a great similarity between the results in the bar graph of protein supply from vegetables and the bar graph of consumption of vegetables. As in consumption, the Asian continent is the continent with the highest protein intake compared to other continents in protein supply from vegetables.



# Visualization of COVID-19 and Food Datasets

In the regression graph, the increase in protein supply from vegetables makes little positive contribution to recovery rates. As in the vegetable consumption graph, since the slope of the line in this regression graph is very low, the strength of the relationship between the two variables remains weak.

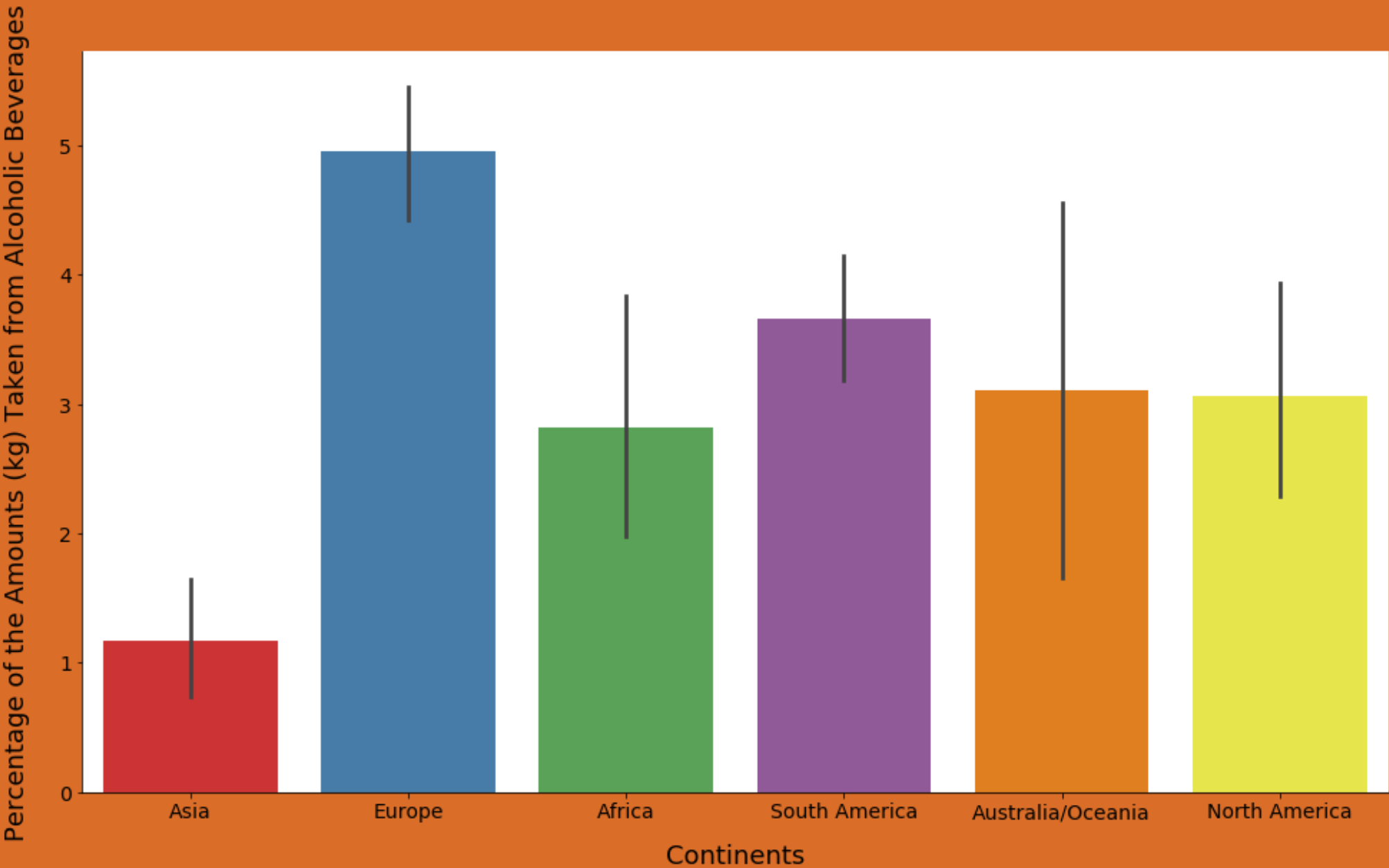
In the other graph , there is a small positive effect on recovery rates as the protein supply from vegetables increases in Asia. One of the reasons for this may be that the Asian continent consumes a lot of vegetables.



# Visualization of COVID-19 and Food Datasets

## 3) Alcoholic Beverages

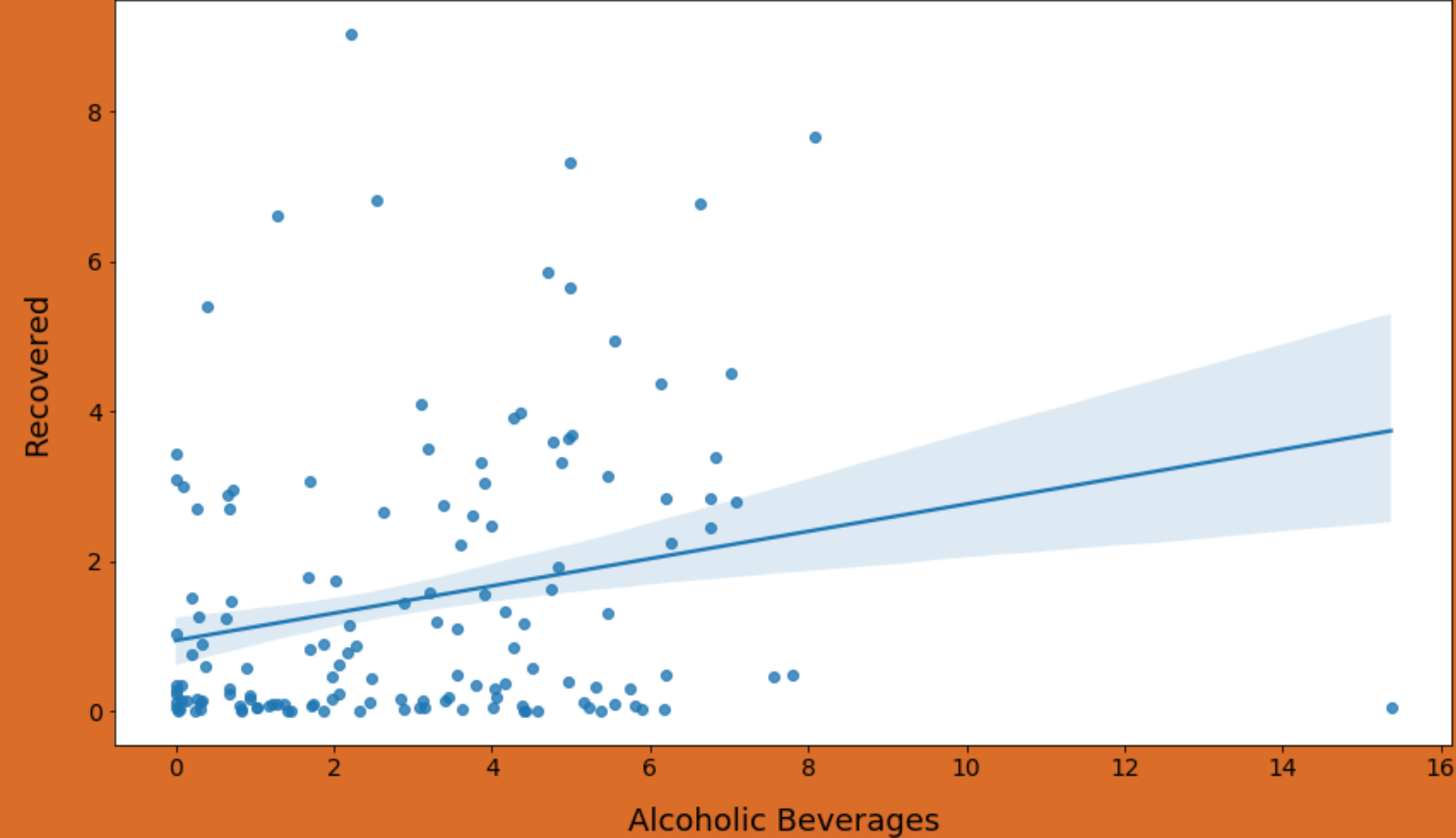
Europe is one of the continents that consumes the most alcoholic beverages in the category of alcoholic beverages. It is understood that the continent that consumes the most after Europe is South America.





# Visualization of COVID-19 and Food Datasets

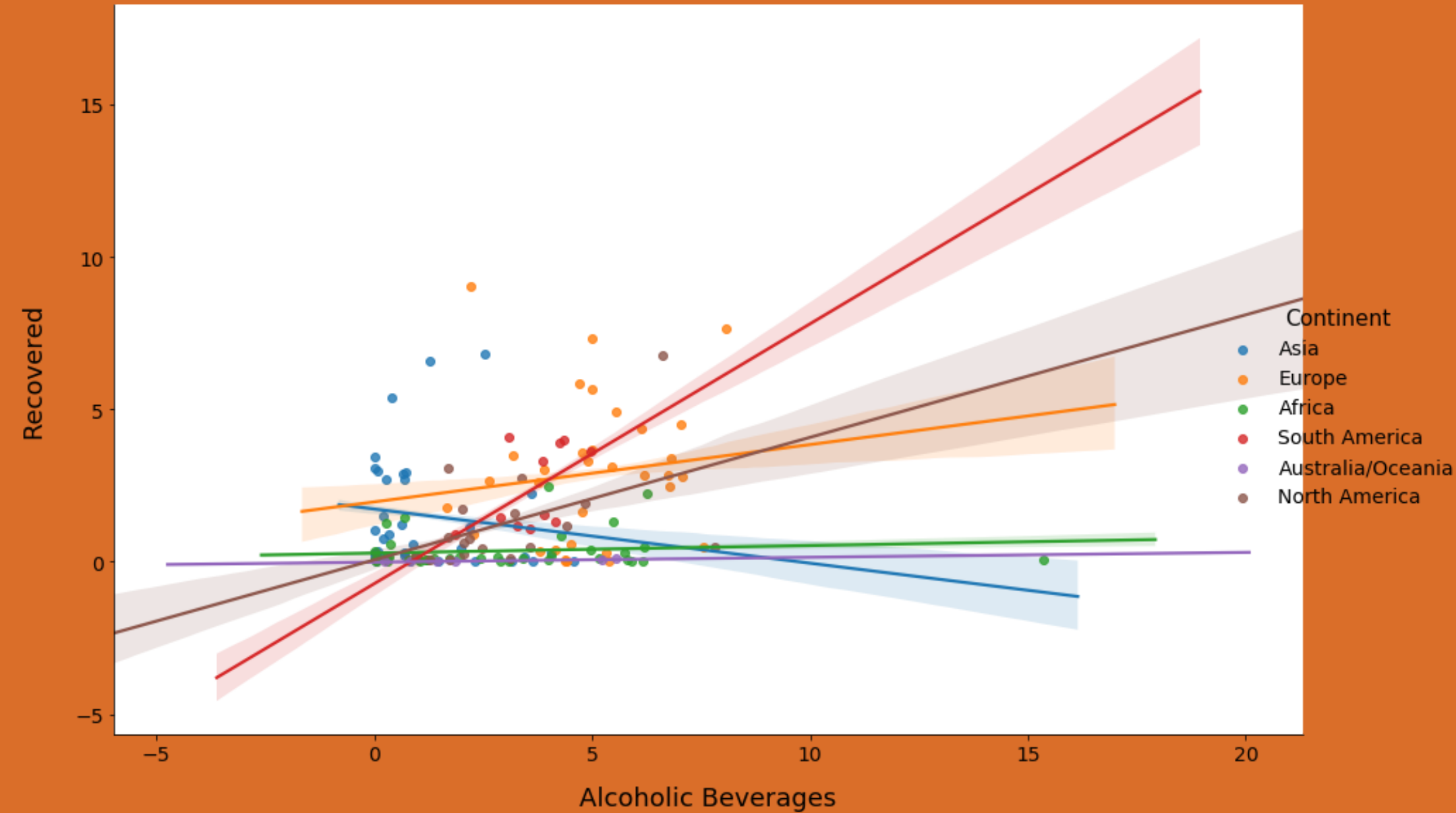
The graph is a regression graph showing the relationship between consumption of alcohol beverages and the recovery rate of coronavirus patients. In this graph, increased consumption of alcoholic beverages leads to increased recovery rates. Although the slope of the regression line in the graph is not high, this result emerges.



*The Relationship Between Recovered & Amounts (kg) Taken from Alcoholic Beverages*

# Visualization of COVID-19 and Food Datasets

When the relationship between recovery rates and alcoholic beverage consumption is analyzed on the basis of continents, the slope of the regression line, especially for South America, is very high. In this case, the increase in the consumption of alcoholic beverages in this continent greatly contributes to the recovery rates.



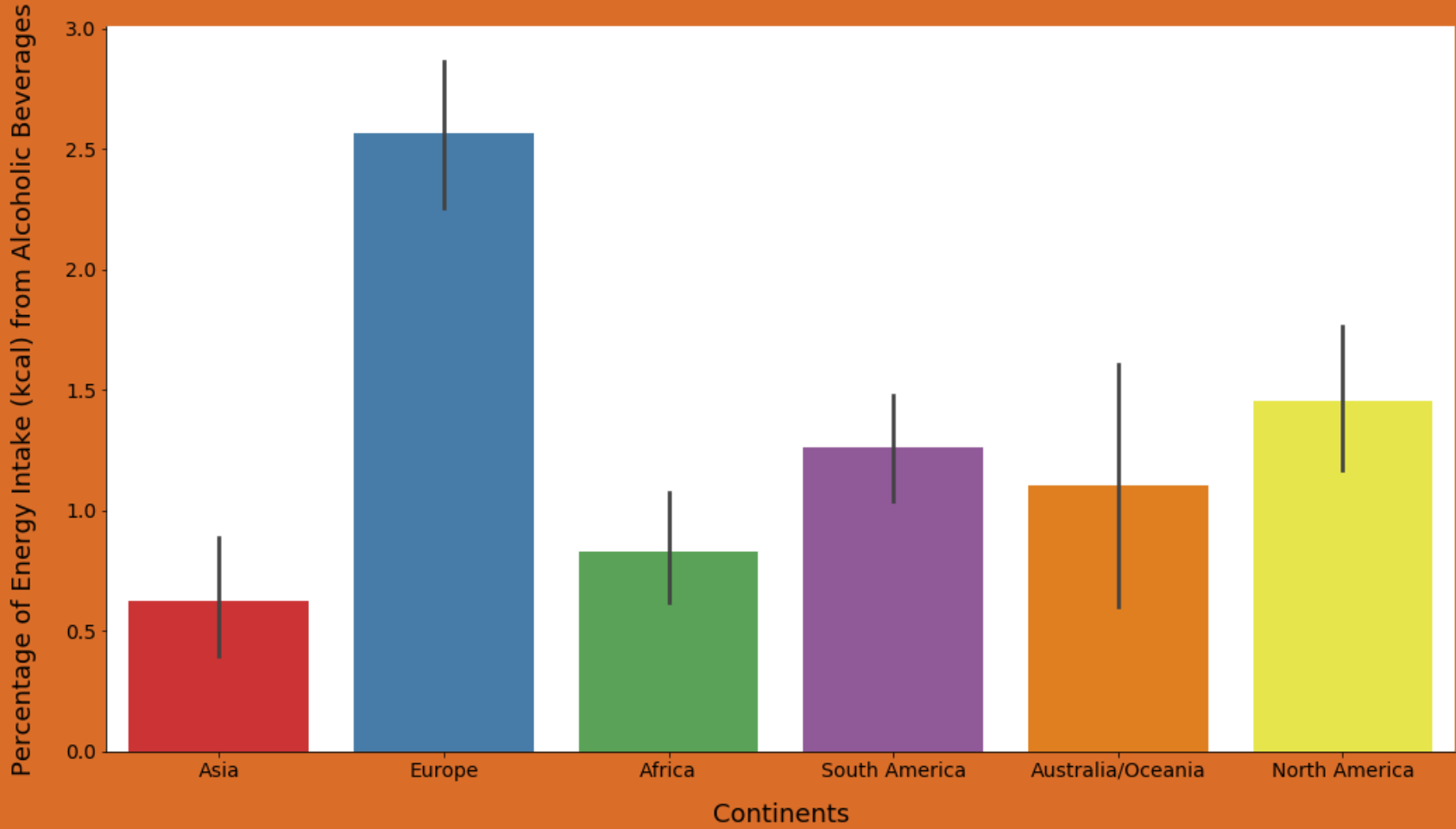
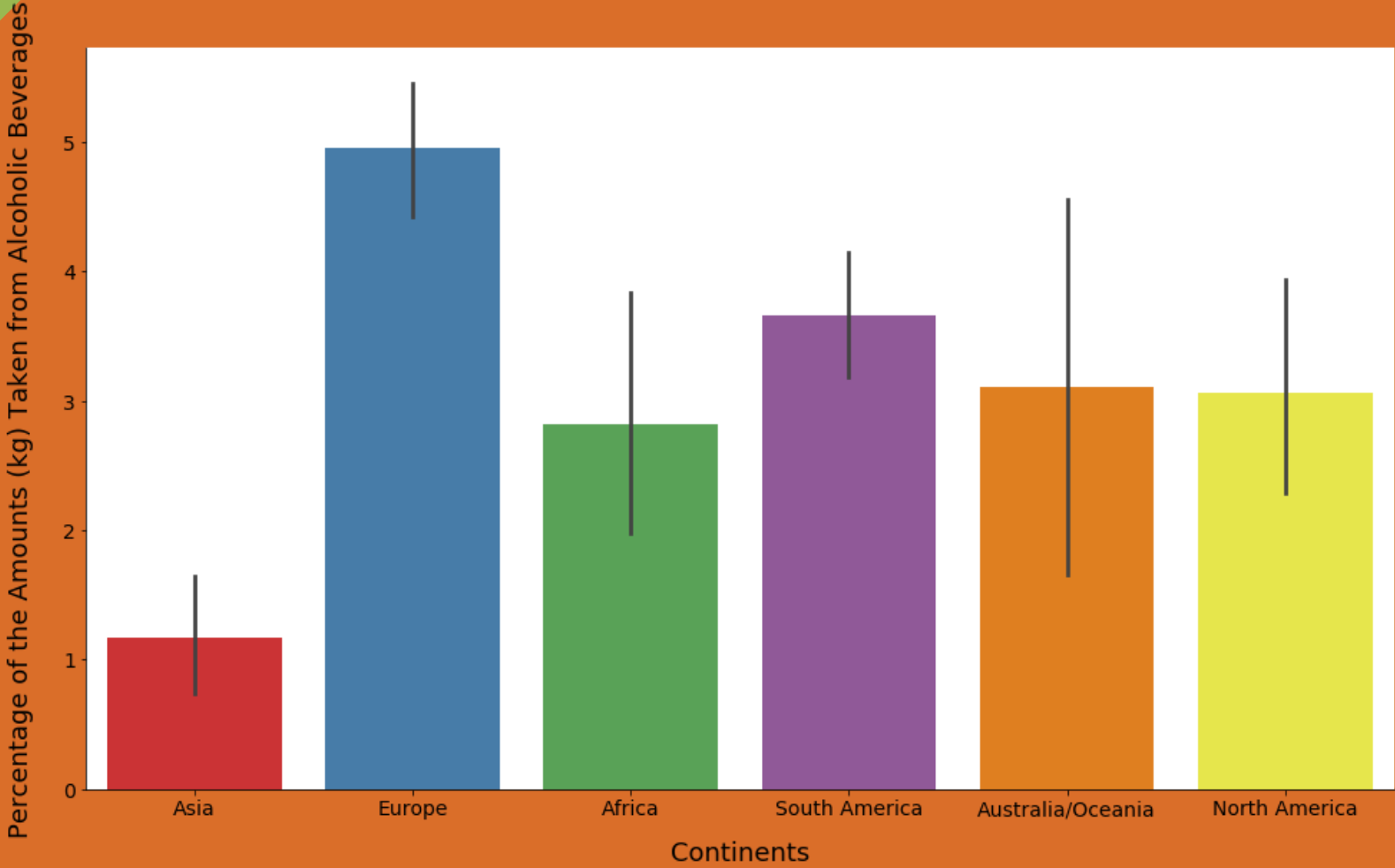
*The Relationship Between Recovered & Amounts (kg) Taken from Alcoholic Beverages by Continents*

# Visualization of COVID-19 and Food Datasets

## 3a) Percentage of Energy Intake (kcal) from Alcoholic Beverages

Energy rates from alcoholic beverages produce similar results as in alcoholic beverage consumption. The European continent, which is the leader in alcoholic beverage consumption, is also the continent that gets the most energy from alcoholic beverages.

The Asian continent, which is in the last place in the consumption of alcoholic beverages, is also in the last place in the energy intake from alcoholic beverages.

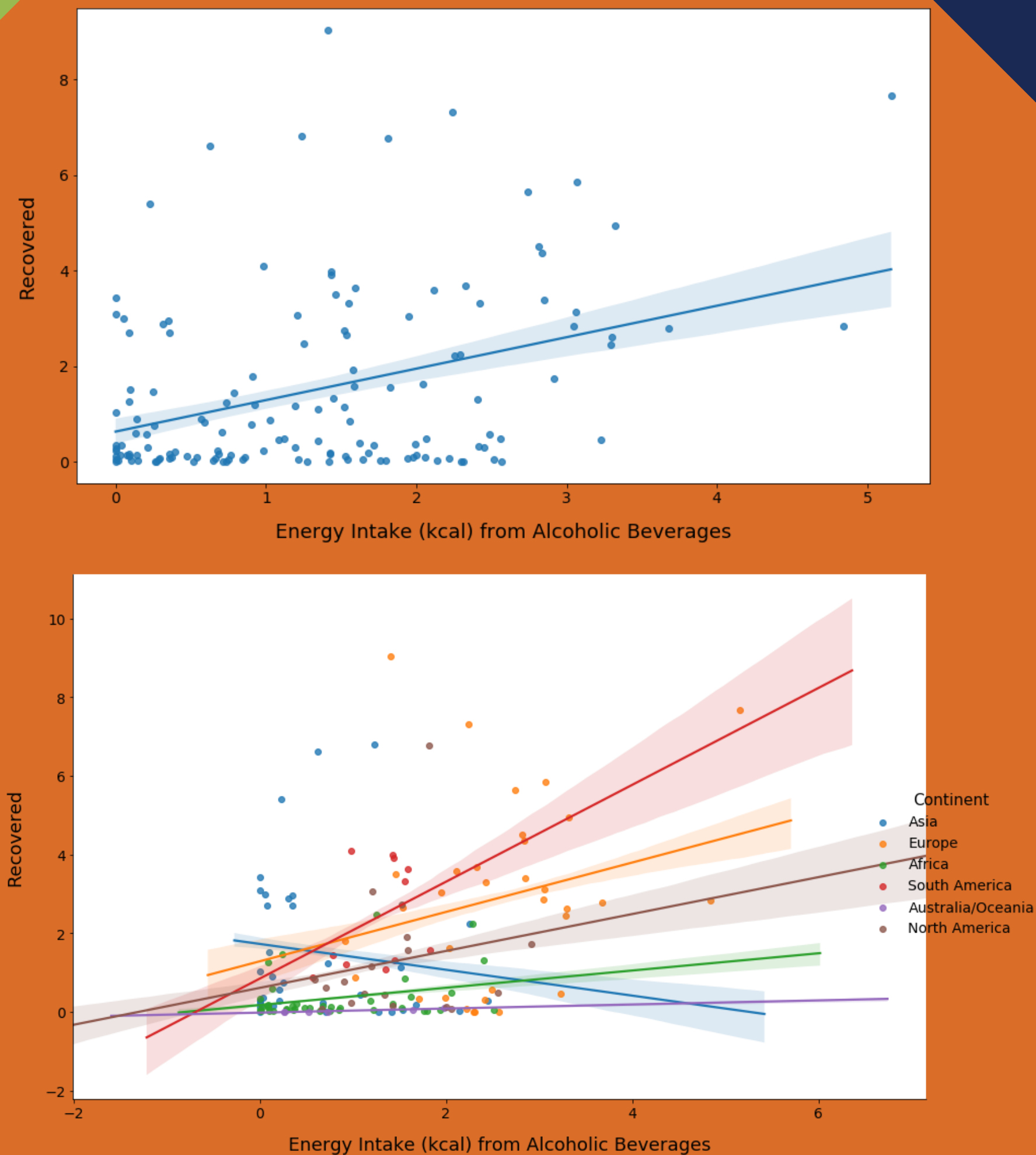




# Visualization of COVID-19 and Food Datasets

The regression graph showing the relationship between the two variables shows that the increase in energy intake from alcoholic beverages has a positive effect on recovery rates.

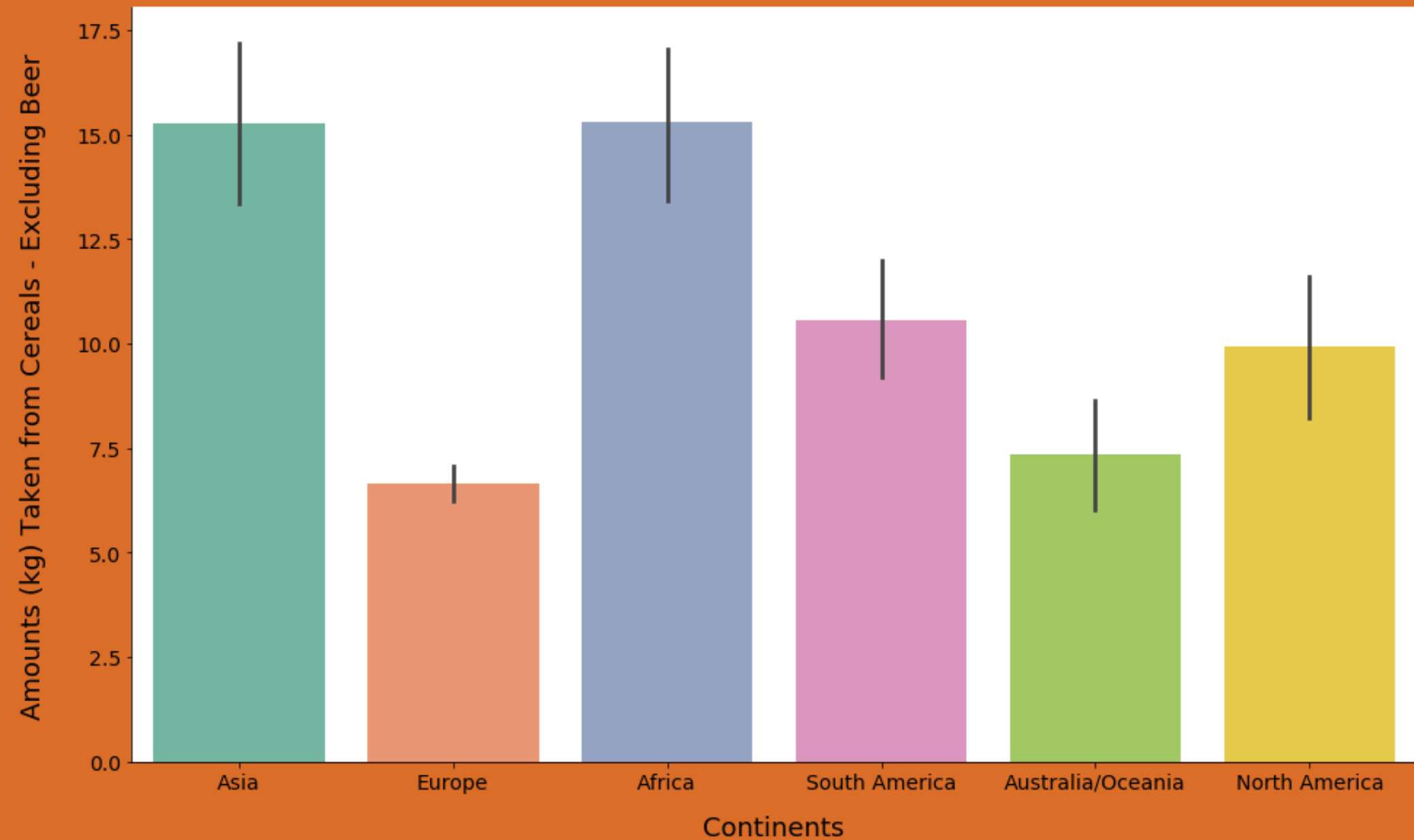
In the regression graph of the continents, when the amount of energy taken from alcoholic beverages increases, recovery rates also increase in South America, North America and Europe. The increase in Africa is less than in other continents because the slope of the regression graph belonging to the African continent is closer to the zero point.



# Visualization of COVID-19 and Food Datasets

## 4) Cereals - Excluding Beer

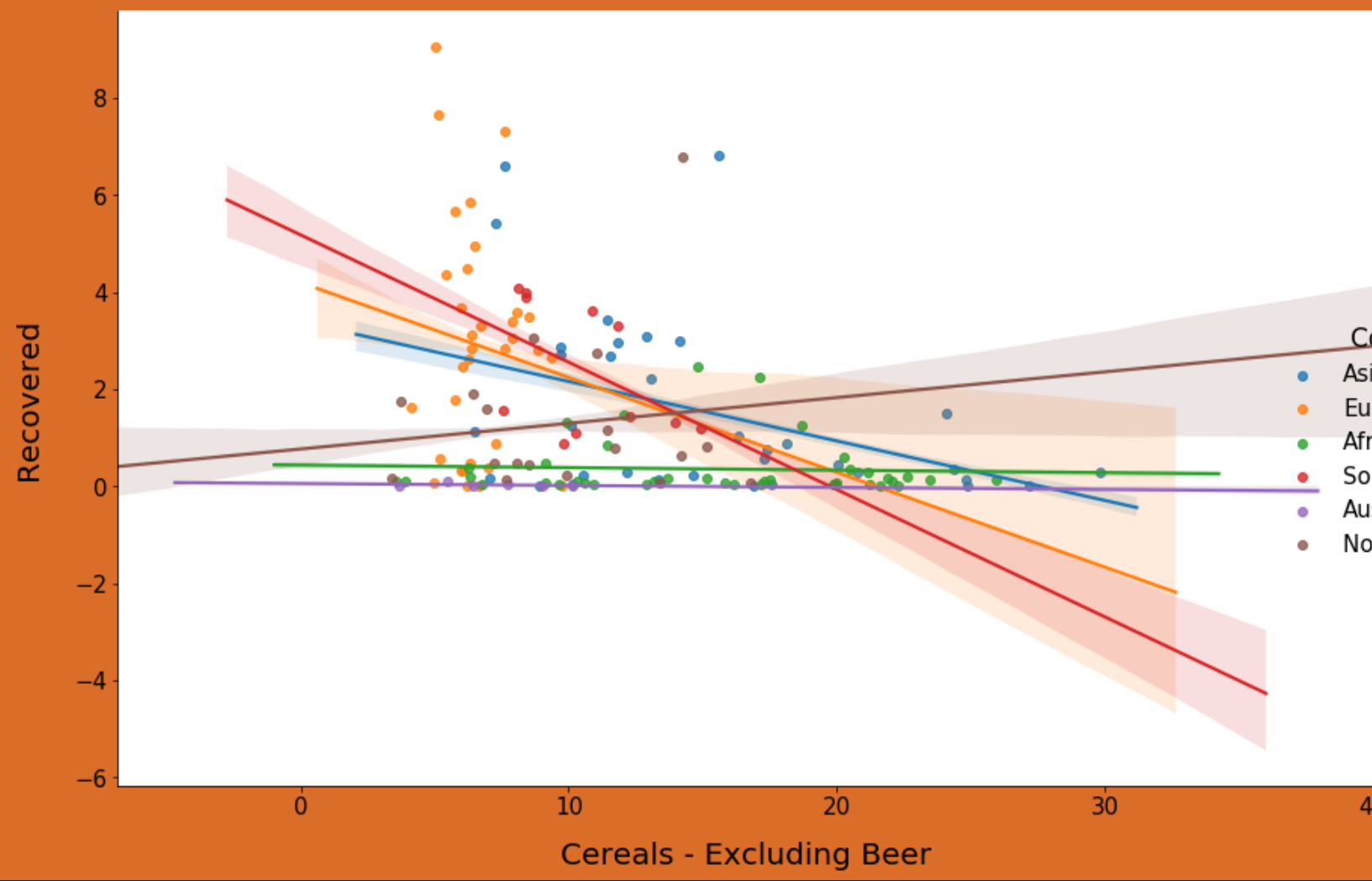
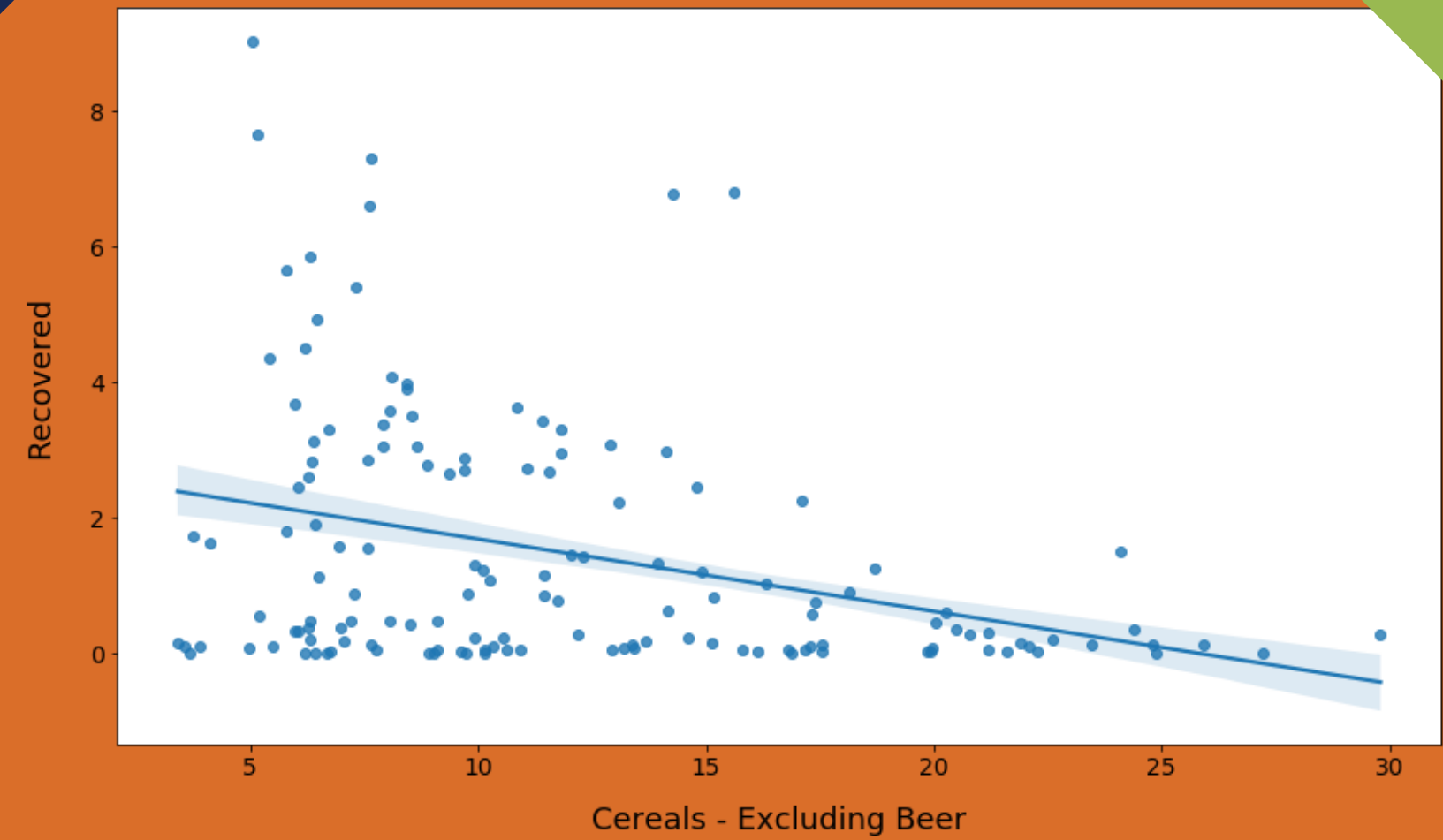
The continents in the first place in cereals consumption are Asia and Africa. Asian and African continents have almost the same consumption rates. The continent with the lowest cereals consumption is Europe.



# Visualization of COVID-19 and Food Datasets

When cereals consumption in the continents increases, the recovery rates of patients with coronavirus are negatively affected. In short, recovery rates are declining.

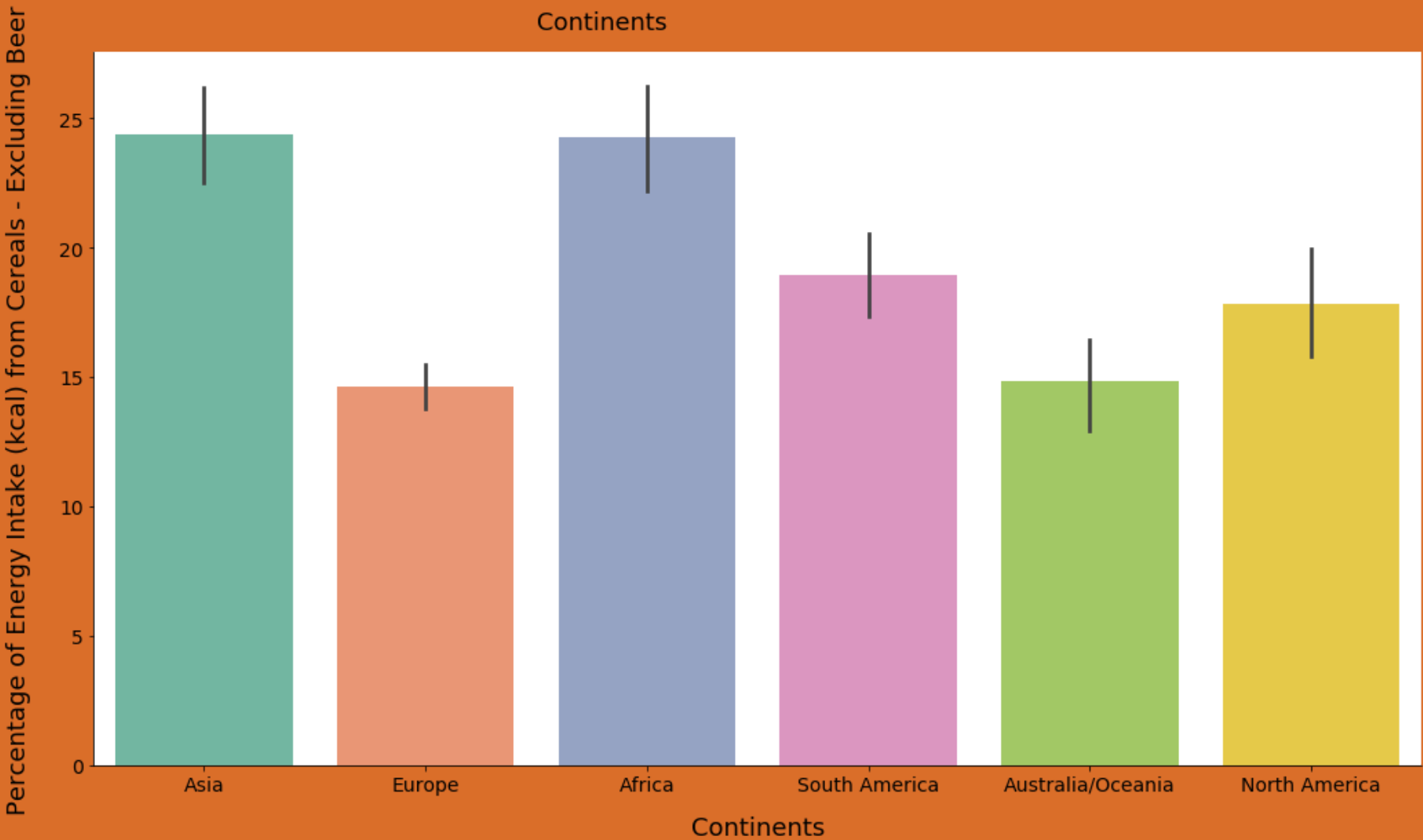
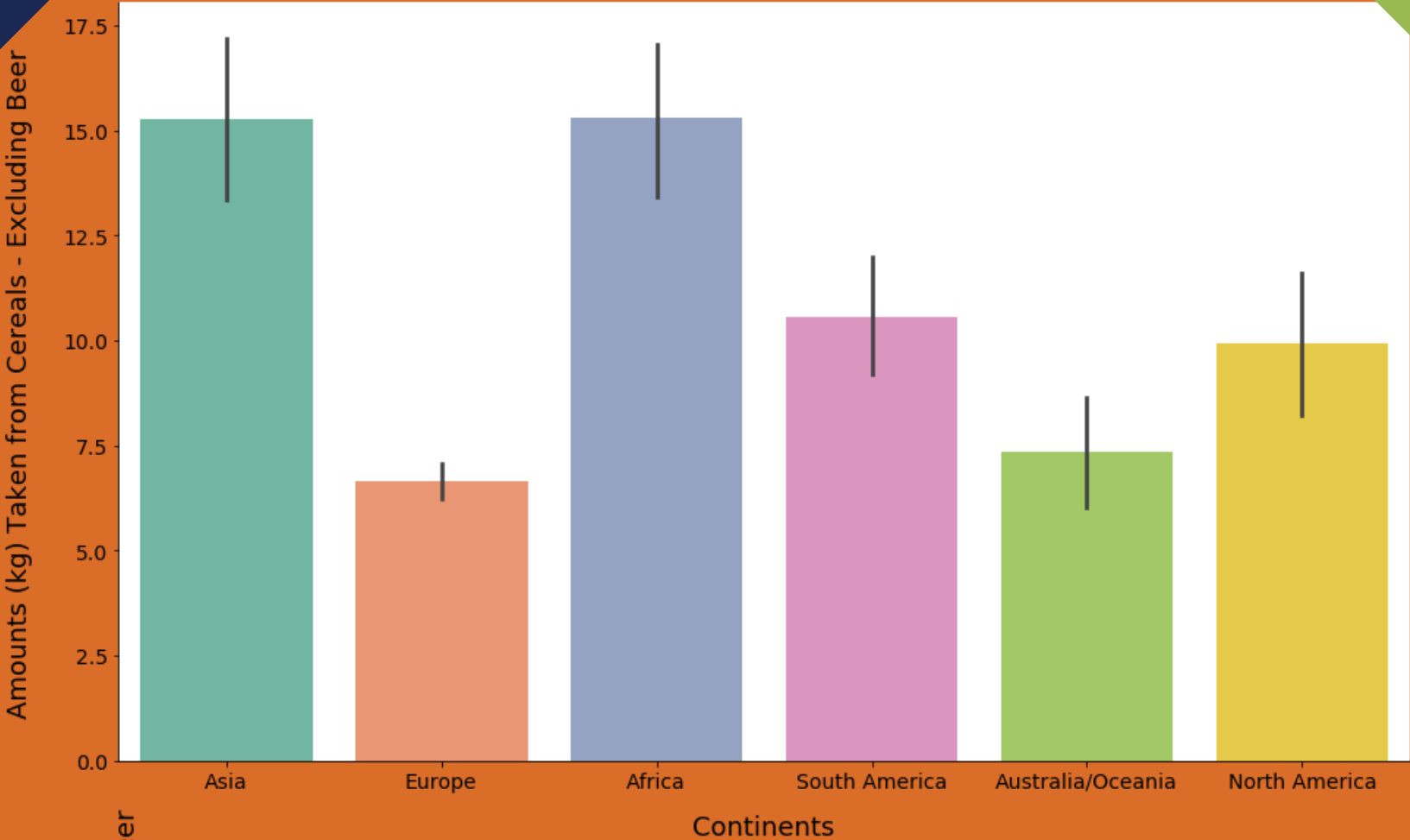
The increase in cereals consumption in South America reduces the recovery rates to lower levels. The slope of the regression line for South America shows how strong the relationship between these two variables is. The situation in South America also exists in continents such as Asia and Europe.



# Visualization of COVID-19 and Food Datasets

4a) Percentage of Energy Intake (kcal) from Cereals - Excluding Beer

Asian and African continents have the highest energy intake from cereals. After Asia and Africa continents, South America is the continent with the highest energy intake from cereals.

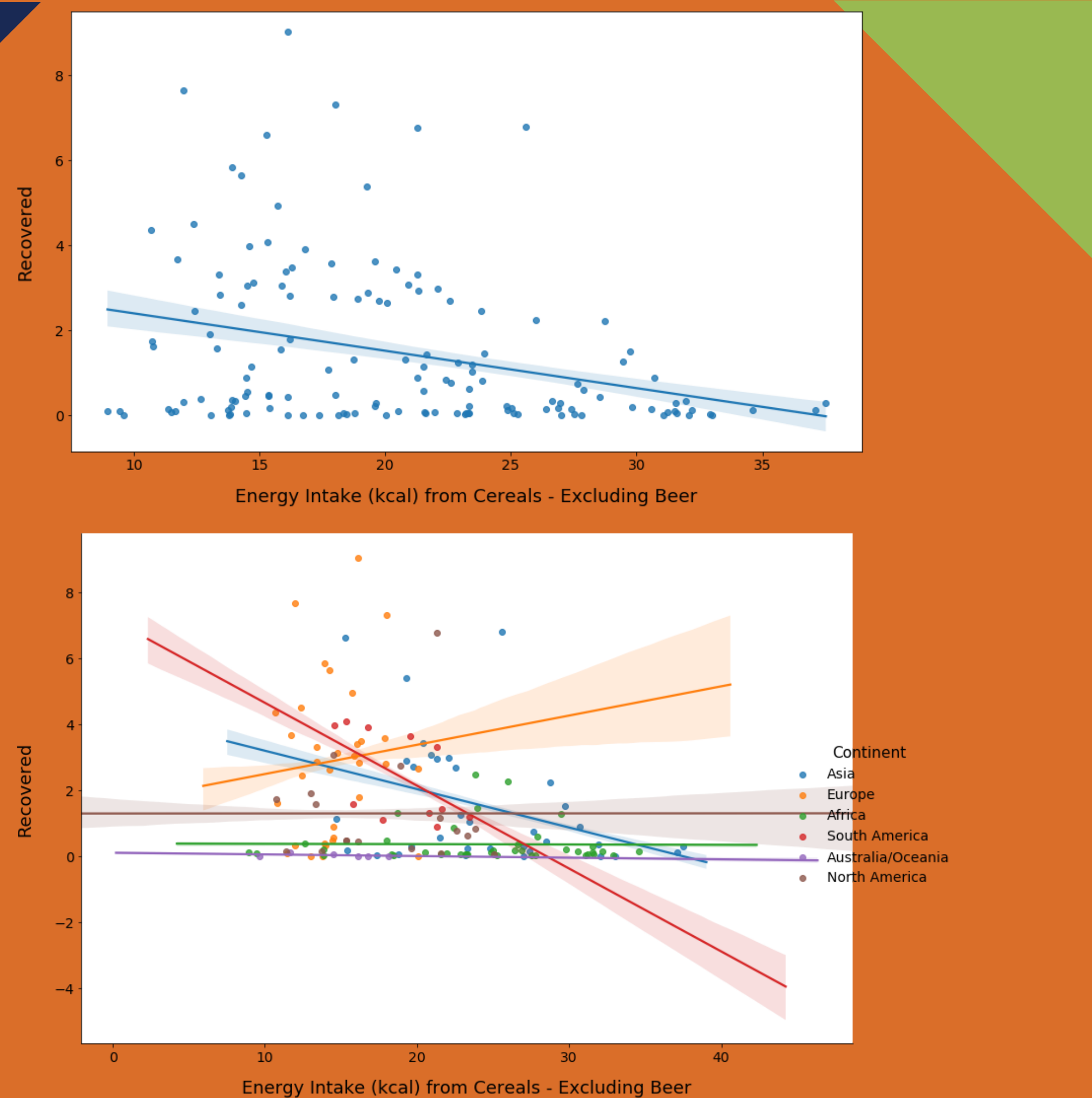




# Visualization of COVID-19 and Food Datasets

As the amount of energy taken from cereals increases, the recovery rate of coronavirus patients decreases.

Due to the increase in the amount of energy taken from cereals, the continent that experienced the sharpest decline in recovery rates is South America. The slope of the regression line for South America gives this result.



# Conclusion of Data Visualization Project

- Animal products, vegetables, and alcoholic beverages increase recovery rates.
- Only the cereals products group could not contribute to the recovery.
- At the same time, consuming too much of a product group or getting too much protein and energy from that product group does not have a large positive relationship with recovery rates.
- Another result is that although a continent consumes a product group more, the amount of protein or energy it receives is less than the other continent.

