Names:	Sean Carney / Cameron Carton
Student Numbers:	19349341 / 19720959
Dataset description:	[Cost of Living by City 2018 Dataset](#).
	This dataset contains information pertaining to the cost of living in various cities around the world. The Indexes in the data show the relative Cost of Living(Plus Rent), Groceries Prices, Restaurant Prices, Rent, and the Local Purchasing Power of each city. The indexes are measured relative to New York City, USA which has an index of 100.00 in every category.
Initial questions:	Which continent has the highest cost of living on average?, How does the cost of living vary within the continents?

Description:	This boxplot visualises the cost of living of each continent. Each box represents the inter-quartile range of the cost of living index in all the cities of the given continent. The white line within the boxes marks the median value of the cost of living index. The lines or 'whiskers' above and below the boxes extend to the maximum and minimum value, respectively, excluding outlier values. The points above and below the whiskers are outliers in the data, the city names can be seen by hovering over the points. Hovering over a box will reveal some of the useful data values.
Insight:	The continent with the highest cost of living and the least variance is Oceania, this is due to the smaller sample size of 15 cities in only 2 countries; Australia and New Zealand, which are both highly-developed countries. If the smaller cities of the island nations of polynesia like Fiji and Samoa were included the data would be very different. For example, Apia, the capital city of Samoa has a [Cost of Living Index of 63.56](#) as of 2022.
	The continent with the greatest disparity in the cost of living between cities is Europe. Europe has both the largest inter-quartile range and the greatest difference in cost of living index between two cities. The poorer Eastern European countries such as Ukraine and Bosnia and Herzegovina are far cheaper to live in than the scandinavian countries and the UK and Ireland, which are all among the most [highly-developed in the world.](#)
	Asia looks to be the continent with the lowest cost of living, as it has a median of just over 41 whereas Africa's is 45.8. The spread of Asia's cost of living is greater than Africa's though and there are multiple outliers above the box. We can see that there are more Asian cities that are cheaper to live in than African cities but the most expensive cities in Asia are far more expensive than those of Africa.
	Another point worth mentioning is the outlier of Hamilton, Bermuda in North America. This extreme outlier is mostly caused by the incredibly high import taxes on goods in Bermuda, which is an island nation and so, needs to import a lot of goods leading to a very high cost of living.
Design considerations:	A boxplot was used as it is a lot easier to see the averages and spread of the quantitative data within each continent. Each box spans the range in which 50% of the cities are located, giving the reader a good idea of the average cost of living in a given continent. The vertical difference from each box is very easy to understand and take away some information quickly. The extent of the whiskers was decreased as it better visualises the range of costs while excluding outliers.
	A [histogram](#) was considered to show the median and deviation of the cost of living within the continents but a boxplot shows this information more succinctly and intuitively.
	The x-axis labels were kept without rotation for easier reading as there was enough space between each label.
	For the colours of the boxes, we looked for [typical colours](#) used to display each continent without confusion. There is [no consensus](#) on the best colour to represent each continent so one of the popular configurations was chosen.
Data filtering and transformation:	We added a new column to the dataset which stores the continent that each city is located in using Python