Global Obesity Analysis

Introduction

What is obesity?

Obesity is a complex disease characterized by an abnormally high level of body fat. Body fat is not an illness in and of itself, but having too much additional fat in your body can alter how it operates. Obesity is generally defined by the Body Mass Index (BMI). BMI is a measure of body fat based on height and weight that is calculated by dividing the body mass (in kilograms) by the square of the body height (in meters). Obesity is often associated with a BMI of 30 or above, according to healthcare professionals. But as BMI does not directly measure body fat, some individuals, such as athletes, may have a BMI in the obese range while having no extra body fat.

What causes obesity?

Many factors impact body weight, including genes, early life effects, poor diets, insufficient physical exercise, and inadequate sleep. Obesity develops when you consume more calories than you burn via typical everyday activity and exercise, despite genetic and hormonal variables. These extra calories are stored as fat by your body. Unhealthy diets - A calorie-dense diet high in fast food and low in fruits and vegetables, as well as a high intake of high-calorie liquids such as alcohol,sugared soft drinks, and large portions, promotes weight gain. Inactivity - Sedentary lifestyles can easily lead to consuming more calories than are burnt via exercise and everyday activities. Staring at computers, tablets and phone screens are examples of sedentary behaviors.

What are some health side effects of obesity?

Raised BMI is a major risk factor for many diseases:

- Cardiovascular diseases (mainly heart disease and stroke)
- High blood pressure
- Diabetes
- Sleep apnea
- Musculoskeletal disorders
- Cancers(endometrial, breast, ovarian, prostate, liver, gallbladder, kidney, and colon)
- Pregnancy problem

Obesity has non-medical side effects as well. These are the side effects that decrease the quality of life for the person.

- Depression
- Disability

- Shame/Guilt
- Social Isolation
- Lower work achievement

These and more are examples of things people can experience when dealing with obesity (Mayo Clinic).

Trends of Obesity Globally

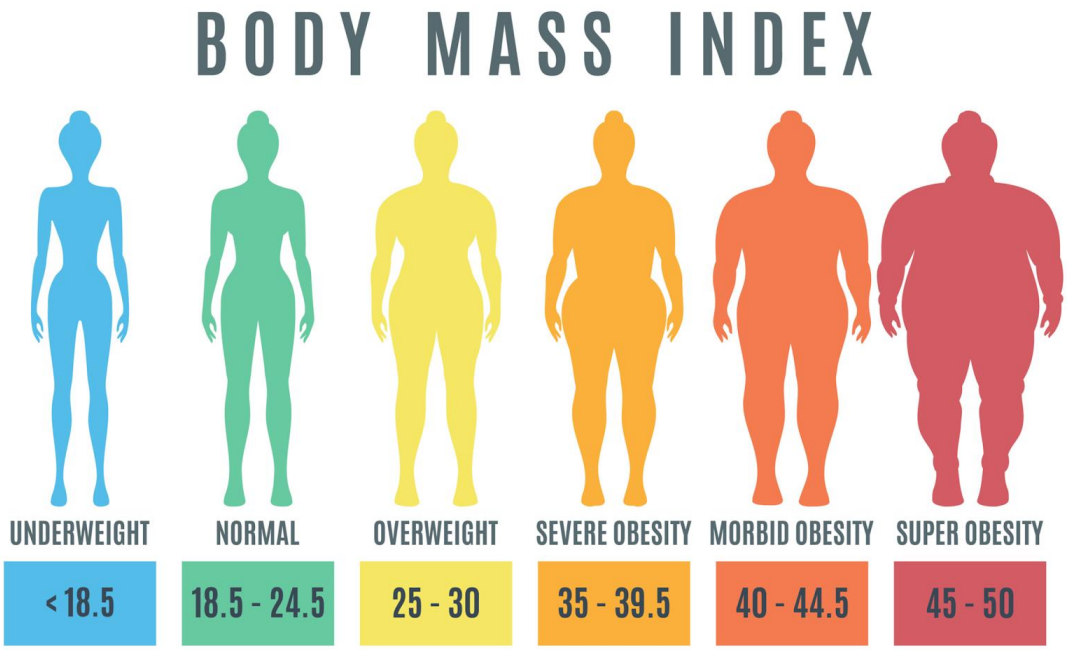
Obesity is a pressing global issue, and it is only getting worse with time. The World Health Organization has described obesity as a global epidemic. The obesity rate has nearly tripled since 1975 over the world, with approximately 13% of adults being obese and approximately 39% of adults being overweight.

Analysis

How We Got the Data:

The World Health Organization's website has a tab leading to the Global Health Observatory. Here there are lists of data sets about countries and health topics. As our group's main focus has been centered around obesity on a global scale, we have been looking into data sets that provide insight to different aspects of the topic. So far these data sets include:

- Prevalence of obesity among adults, BMI & GreaterEqual; 30 (age-standardized estimate)(%)
- Prevalence of overweight among adults, BMI & Greater Equal; 25 (age-standardized estimate)(%)
- Mean BMI (kg/m2) (age-standardized estimate)
- Has conducted a recent, national adult risk factor survey covering overweight and obesity
- Has conducted a recent, national adult risk factor survey covering unhealthy diet
- Has conducted a recent, national adult risk factor survey covering physical inactivity
- Probability of Deaths
- Raised Blood Pressure



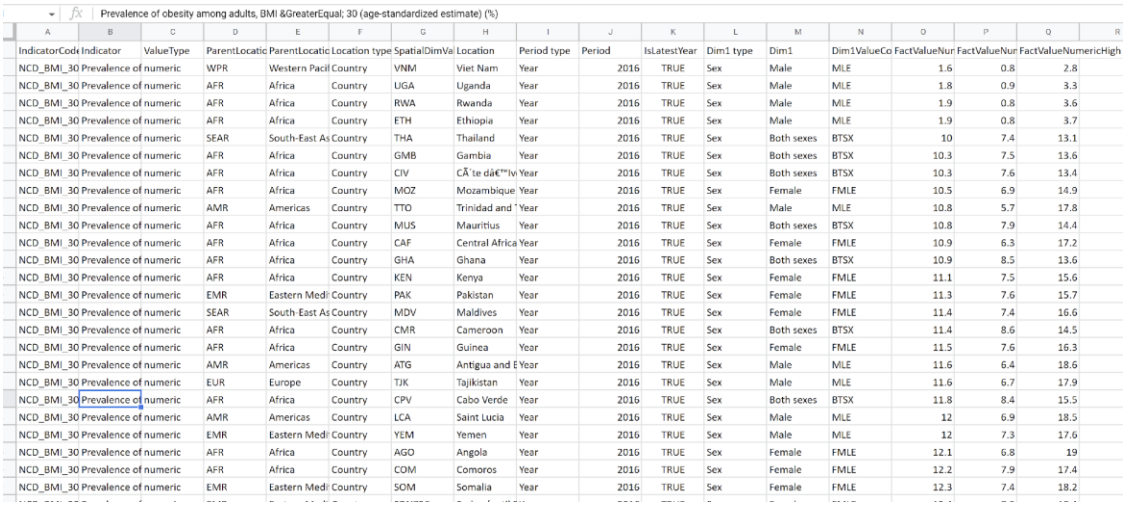
https://www.lipedema.net/lipedema-lymphedema-obesity.html

Data Cleaning and Preparation:

In order to clean and prepare the datasets, we performed the following steps:

- Handling the null values: For all the 6 datasets, we analaysed the missing data values in key variables. For example, for the Prevalence of Obesity & Overweight dataset,the null values accounted for only 0.01% of the data, hence we decided to remove them from the dataset.
- Renaming and selecting relevant columns: After doing our initial analysis, we selected the columns with relevant information and renamed those columns to make them more intuitive
- Merging similar datasets: We realised that some of our datasets were correlated and had similar data structure. So we combined such datasets into one making it easier to use them. For example, Prevalence of Obesity and Prevalence of Overweight had the same data structure, so we merged these two datasets into one (refer to the screenshots below for more details). The same technique was applied for “Risk factor survey covering overweight and obesity”, “Risk factor survey covering unhealthy diet” and “Risk factor survey covering physical inactivity”.
- Analyzing and treating outliers: We created boxplots to understand the data outliers. We decided to leave these outliers in because it was clear that these dat points were not due to measurement issues and had actual value.
- Appropriate data type conversion: We converted the data type of few columns into the required data type, wherever needed.

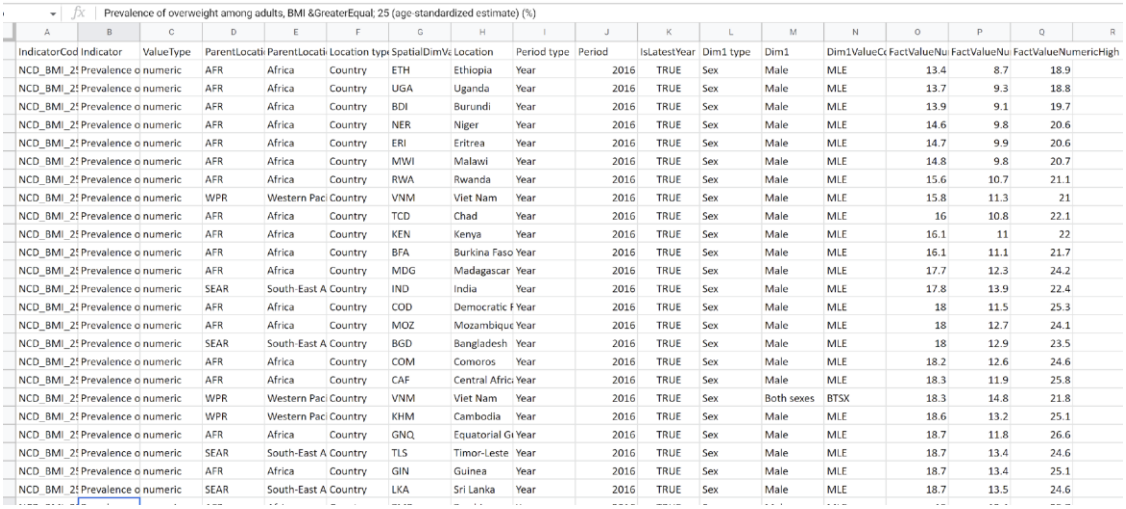
Here are some “before” and “after” screenshots of our data to give you an idea of our data sets:



Prevalence of obesity among adults, BMI ≥ 30 (age-standardized estimate) (%)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	IndicatorCod	Indicator	ValueType	ParentLocatio	ParentLocatio	Location type	SpatialDimVa	Location	Period type	Period	IsLatestYear	Dim1 type	Dim1	Dim1ValueCo	FactValueNur	FactValueNur	FactValueNumericHigh	
	NCD_BMI_30	Prevalence of	numeric	WPR	Western Pacil	Country	VNM	Viet Nam	Year	2016	TRUE	Sex	Male	MLE	1.6	0.8	2.8	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	UGA	Uganda	Year	2016	TRUE	Sex	Male	MLE	1.8	0.9	3.3	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	RWA	Rwanda	Year	2016	TRUE	Sex	Male	MLE	1.9	0.8	3.6	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	ETH	Ethiopia	Year	2016	TRUE	Sex	Male	MLE	1.9	0.8	3.7	
	NCD_BMI_30	Prevalence of	numeric	SEAR	South-East As	Country	THA	Thailand	Year	2016	TRUE	Sex	Both sexes	BTSX	10	7.4	13.1	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	GMB	Gambia	Year	2016	TRUE	Sex	Both sexes	BTSX	10.3	7.5	13.6	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	CIV	CÃ´te dâ€™Iv	Year	2016	TRUE	Sex	Both sexes	BTSX	10.3	7.6	13.4	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	MOZ	Mozambique	Year	2016	TRUE	Sex	Female	FMLE	10.5	6.9	14.9	
	NCD_BMI_30	Prevalence of	numeric	AMR	Americas	Country	TTO	Trinidad and '	Year	2016	TRUE	Sex	Male	MLE	10.8	5.7	17.8	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	MUS	Mauritius	Year	2016	TRUE	Sex	Both sexes	BTSX	10.8	7.9	14.4	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	CAF	Central Africa	Year	2016	TRUE	Sex	Female	FMLE	10.9	6.3	17.2	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	GHA	Ghana	Year	2016	TRUE	Sex	Both sexes	BTSX	10.9	8.5	13.6	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	KEN	Kenya	Year	2016	TRUE	Sex	Female	FMLE	11.1	7.5	15.6	
	NCD_BMI_30	Prevalence of	numeric	EMR	Eastern Medi	Country	PAK	Pakistan	Year	2016	TRUE	Sex	Female	FMLE	11.3	7.6	15.7	
	NCD_BMI_30	Prevalence of	numeric	SEAR	South-East As	Country	MDV	Maldives	Year	2016	TRUE	Sex	Female	FMLE	11.4	7.4	16.6	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	CMR	Cameroon	Year	2016	TRUE	Sex	Both sexes	BTSX	11.4	8.6	14.5	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	GIN	Guinea	Year	2016	TRUE	Sex	Female	FMLE	11.5	7.6	16.3	
	NCD_BMI_30	Prevalence of	numeric	AMR	Americas	Country	ATG	Antigua and E	Year	2016	TRUE	Sex	Male	MLE	11.6	6.4	18.6	
	NCD_BMI_30	Prevalence of	numeric	EUR	Europe	Country	TJK	Tajikistan	Year	2016	TRUE	Sex	Male	MLE	11.6	6.7	17.9	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	CPV	Cabo Verde	Year	2016	TRUE	Sex	Both sexes	BTSX	11.8	8.4	15.5	
	NCD_BMI_30	Prevalence of	numeric	AMR	Americas	Country	LCA	Saint Lucia	Year	2016	TRUE	Sex	Male	MLE	12	6.9	18.5	
	NCD_BMI_30	Prevalence of	numeric	EMR	Eastern Medi	Country	YEM	Yemen	Year	2016	TRUE	Sex	Male	MLE	12	7.3	17.6	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	AGO	Angola	Year	2016	TRUE	Sex	Female	FMLE	12.1	6.8	19	
	NCD_BMI_30	Prevalence of	numeric	AFR	Africa	Country	COM	Comoros	Year	2016	TRUE	Sex	Female	FMLE	12.2	7.9	17.4	
	NCD_BMI_30	Prevalence of	numeric	EMR	Eastern Medi	Country	SOM	Somalia	Year	2016	TRUE	Sex	Female	FMLE	12.3	7.4	18.2	

Prevalence of Obese Before



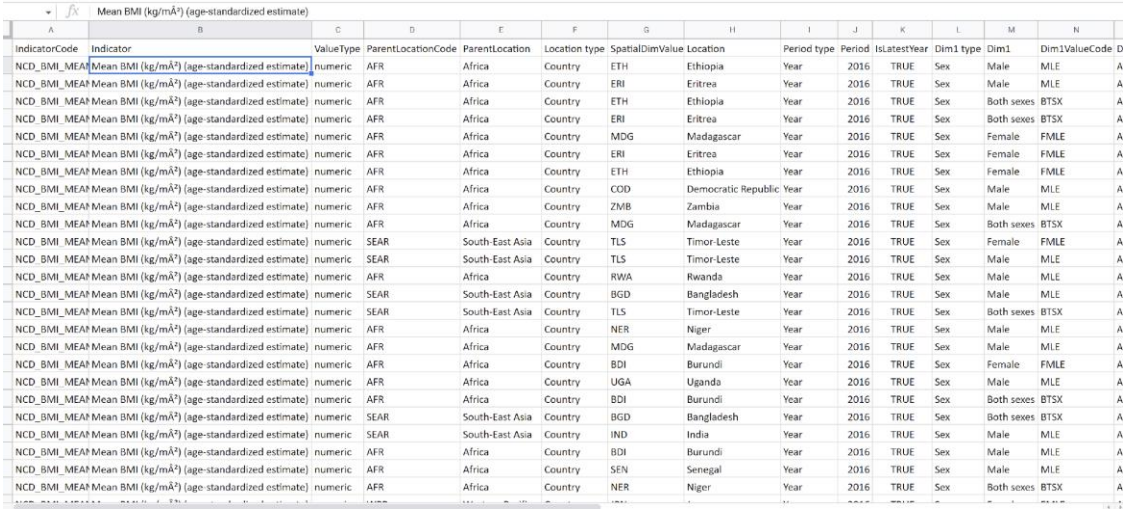
Prevalence of overweight among adults, BMI ≥ 25 (age-standardized estimate) (%)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	IndicatorCod	Indicator	ValueType	ParentLocatio	ParentLocatio	Location type	SpatialDimVa	Location	Period type	Period	IsLatestYear	Dim1 type	Dim1	Dim1ValueCo	FactValueNu	FactValueNu	FactValueNumericHigh	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	ETH	Ethiopia	Year	2016	TRUE	Sex	Male	MLE	13.4	8.7	18.9	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	UGA	Uganda	Year	2016	TRUE	Sex	Male	MLE	13.7	9.3	18.8	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	BDI	Burundi	Year	2016	TRUE	Sex	Male	MLE	13.9	9.1	19.7	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	NER	Niger	Year	2016	TRUE	Sex	Male	MLE	14.6	9.8	20.6	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	ERI	Eritrea	Year	2016	TRUE	Sex	Male	MLE	14.7	9.9	20.6	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	MWI	Malawi	Year	2016	TRUE	Sex	Male	MLE	14.8	9.8	20.7	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	RWA	Rwanda	Year	2016	TRUE	Sex	Male	MLE	15.6	10.7	21.1	
	NCD_BMI_25	Prevalence o	numeric	WPR	Western Paci	Country	VNM	Viet Nam	Year	2016	TRUE	Sex	Male	MLE	15.8	11.3	21	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	TCD	Chad	Year	2016	TRUE	Sex	Male	MLE	16	10.8	22.1	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	KEN	Kenya	Year	2016	TRUE	Sex	Male	MLE	16.1	11	22	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	BFA	Burkina Faso	Year	2016	TRUE	Sex	Male	MLE	16.1	11.1	21.7	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	MDG	Madagascar	Year	2016	TRUE	Sex	Male	MLE	17.7	12.3	24.2	
	NCD_BMI_25	Prevalence o	numeric	SEAR	South-East A	Country	IND	India	Year	2016	TRUE	Sex	Male	MLE	17.8	13.9	22.4	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	COD	Democratic F	Year	2016	TRUE	Sex	Male	MLE	18	11.5	25.3	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	MOZ	Mozambique	Year	2016	TRUE	Sex	Male	MLE	18	12.7	24.1	
	NCD_BMI_25	Prevalence o	numeric	SEAR	South-East A	Country	BGD	Bangladesh	Year	2016	TRUE	Sex	Male	MLE	18	12.9	23.5	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	COM	Comoros	Year	2016	TRUE	Sex	Male	MLE	18.2	12.6	24.6	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	CAF	Central Afric	Year	2016	TRUE	Sex	Male	MLE	18.3	11.9	25.8	
	NCD_BMI_25	Prevalence o	numeric	WPR	Western Paci	Country	VNM	Viet Nam	Year	2016	TRUE	Sex	Both sexes	BTSX	18.3	14.8	21.8	
	NCD_BMI_25	Prevalence o	numeric	WPR	Western Paci	Country	KHM	Cambodia	Year	2016	TRUE	Sex	Male	MLE	18.6	13.2	25.1	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	GNQ	Equatorial Gi	Year	2016	TRUE	Sex	Male	MLE	18.7	11.8	26.6	
	NCD_BMI_25	Prevalence o	numeric	SEAR	South-East A	Country	TLS	Timor-Leste	Year	2016	TRUE	Sex	Male	MLE	18.7	13.4	24.6	
	NCD_BMI_25	Prevalence o	numeric	AFR	Africa	Country	GIN	Guinea	Year	2016	TRUE	Sex	Male	MLE	18.7	13.4	25.1	
	NCD_BMI_25	Prevalence o	numeric	SEAR	South-East A	Country	LKA	Sri Lanka	Year	2016	TRUE	Sex	Male	MLE	18.7	13.5	24.6	

Prevalence of Overweight Before

	A	B	C	D	E	F	G	H	I
		ParentLocationC	ParentLocation	LocationCode	Period	SexCode	Sex	Prev_overwt	Prev_obesity
1		AFR	Africa	ETH	2016	MLE	Male	13.4	1.9
2		AFR	Africa	UGA	2016	MLE	Male	13.7	1.8
3		AFR	Africa	BDI	2016	MLE	Male	13.9	2.1
4		AFR	Africa	NER	2016	MLE	Male	14.6	2.5
5		AFR	Africa	ERI	2016	MLE	Male	14.7	2
6		AFR	Africa	MWI	2016	MLE	Male	14.8	2.2
7		AFR	Africa	RWA	2016	MLE	Male	15.6	1.9
8		WPR	Western Pacific	VNM	2016	MLE	Male	15.8	1.6
9		AFR	Africa	TCD	2016	MLE	Male	16	3.1
10		AFR	Africa	KEN	2016	MLE	Male	16.1	2.8
11		AFR	Africa	BFA	2016	MLE	Male	16.1	2.6
12		AFR	Africa	MDG	2016	MLE	Male	17.7	3
13		SEAR	South-East Asia	IND	2016	MLE	Male	17.8	2.7
14		AFR	Africa	COD	2016	MLE	Male	18	3.6
15		AFR	Africa	MOZ	2016	MLE	Male	18	3.3
16		SEAR	South-East Asia	BGD	2016	MLE	Male	18	2.3
17		AFR	Africa	COM	2016	MLE	Male	18.2	3.3
18		AFR	Africa	CAF	2016	MLE	Male	18.3	3.7
19		WPR	Western Pacific	VNM	2016	BTSX	Both sexes	18.3	2.1
20		WPR	Western Pacific	KHM	2016	MLE	Male	18.6	2.7
21		AFR	Africa	GNQ	2016	MLE	Male	18.7	3.8
22		SEAR	South-East Asia	TLS	2016	MLE	Male	18.7	2.6
23		AFR	Africa	GIN	2016	MLE	Male	18.7	3.7
24		SEAR	South-East Asia	LKA	2016	MLE	Male	18.7	2.9
25		AFR	Africa	ZMB	2016	MLE	Male	19	3.6
26		AFR	Africa	SLE	2016	MLE	Male	19.1	3.8
27		SEAR	South-East Asia	NPL	2016	MLE	Male	19.1	2.7
28		AFR	Africa	AGO	2016	MLE	Male	19.5	4

Prevalence of Obese and Overweight After



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	IndicatorCode	Indicator	ValueType	ParentLocationCode	ParentLocation	Location type	SpatialDimValue	Location	Period type	Period	IsLatestYear	Dim1 type	Dim1	Dim1ValueCode
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	ETH	Ethiopia	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	ERI	Eritrea	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	ETH	Ethiopia	Year	2016	TRUE	Sex	Both sexes	BTSX
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	ERI	Eritrea	Year	2016	TRUE	Sex	Both sexes	BTSX
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	MDG	Madagascar	Year	2016	TRUE	Sex	Female	FMLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	ERI	Eritrea	Year	2016	TRUE	Sex	Female	FMLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	ETH	Ethiopia	Year	2016	TRUE	Sex	Female	FMLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	COD	Democratic Republic	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	ZMB	Zambia	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	MDG	Madagascar	Year	2016	TRUE	Sex	Both sexes	BTSX
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	SEAR	South-East Asia	Country	TLS	Timor-Leste	Year	2016	TRUE	Sex	Female	FMLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	SEAR	South-East Asia	Country	TLS	Timor-Leste	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	RWA	Rwanda	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	SEAR	South-East Asia	Country	BGD	Bangladesh	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	SEAR	South-East Asia	Country	TLS	Timor-Leste	Year	2016	TRUE	Sex	Both sexes	BTSX
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	NER	Niger	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	MDG	Madagascar	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	BDI	Burundi	Year	2016	TRUE	Sex	Female	FMLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	UGA	Uganda	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	BDI	Burundi	Year	2016	TRUE	Sex	Both sexes	BTSX
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	SEAR	South-East Asia	Country	BGD	Bangladesh	Year	2016	TRUE	Sex	Both sexes	BTSX
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	SEAR	South-East Asia	Country	IND	India	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	BDI	Burundi	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	SEN	Senegal	Year	2016	TRUE	Sex	Male	MLE
	NCD_BMI_MEAN	Mean BMI (kg/mÂ²) (age-standardized estimate)	numeric	AFR	Africa	Country	NER	Niger	Year	2016	TRUE	Sex	Both sexes	BTSX

Mean BMI Before Part 1

Mean BMI (kg/mÂ²) (age-standardized estimate)

M	N	O	P	Q	R	S	T	U
Dim1	Dim1ValueCode	Dim2 type	Dim2	Dim2ValueCode	FactValueNumeric	FactValueNumer	FactValueNumer	Value
Male	MLE	Age Group	18+ years	YEARS18-PLUS	20.1	19	21	20.1 [19 â€“ 21]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	20.2	19.2	21.3	20.2 [19.2 â€“ 21.3]
Both sexes	BTSX	Age Group	18+ years	YEARS18-PLUS	20.6	19.8	21.3	20.6 [19.8 â€“ 21.3]
Both sexes	BTSX	Age Group	18+ years	YEARS18-PLUS	20.7	19.9	21.5	20.7 [19.9 â€“ 21.5]
Female	FMLE	Age Group	18+ years	YEARS18-PLUS	21.1	19.9	22.3	21.1 [19.9 â€“ 22.3]
Female	FMLE	Age Group	18+ years	YEARS18-PLUS	21.1	20	22.2	21.1 [20 â€“ 22.2]
Female	FMLE	Age Group	18+ years	YEARS18-PLUS	21.1	20	22.2	21.1 [20 â€“ 22.2]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.3	19.4	23.3	21.3 [19.4 â€“ 23.3]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.3	19.8	22.9	21.3 [19.8 â€“ 22.9]
Both sexes	BTSX	Age Group	18+ years	YEARS18-PLUS	21.3	20.3	22.3	21.3 [20.3 â€“ 22.3]
Female	FMLE	Age Group	18+ years	YEARS18-PLUS	21.3	20.5	22.1	21.3 [20.5 â€“ 22.1]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.3	20.5	22.1	21.3 [20.5 â€“ 22.1]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.3	20.7	21.8	21.3 [20.7 â€“ 21.8]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.3	20.7	21.9	21.3 [20.7 â€“ 21.9]
Both sexes	BTSX	Age Group	18+ years	YEARS18-PLUS	21.3	20.8	21.9	21.3 [20.8 â€“ 21.9]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.4	20	22.7	21.4 [20 â€“ 22.7]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.5	19.8	23.2	21.5 [19.8 â€“ 23.2]
Female	FMLE	Age Group	18+ years	YEARS18-PLUS	21.6	20.3	22.9	21.6 [20.3 â€“ 22.9]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.6	20.9	22.3	21.6 [20.9 â€“ 22.3]
Both sexes	BTSX	Age Group	18+ years	YEARS18-PLUS	21.7	20.3	23.1	21.7 [20.3 â€“ 23.1]
Both sexes	BTSX	Age Group	18+ years	YEARS18-PLUS	21.7	21.2	22.1	21.7 [21.2 â€“ 22.1]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.7	21.2	22.3	21.7 [21.2 â€“ 22.3]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.8	19.3	24.2	21.8 [19.3 â€“ 24.2]
Male	MLE	Age Group	18+ years	YEARS18-PLUS	21.8	20.8	22.8	21.8 [20.8 â€“ 22.8]
Both sexes	BTSX	Age Group	18+ years	YEARS18-PLUS	21.8	21	22.6	21.8 [21 â€“ 22.6]

Mean BMI Before Part 2

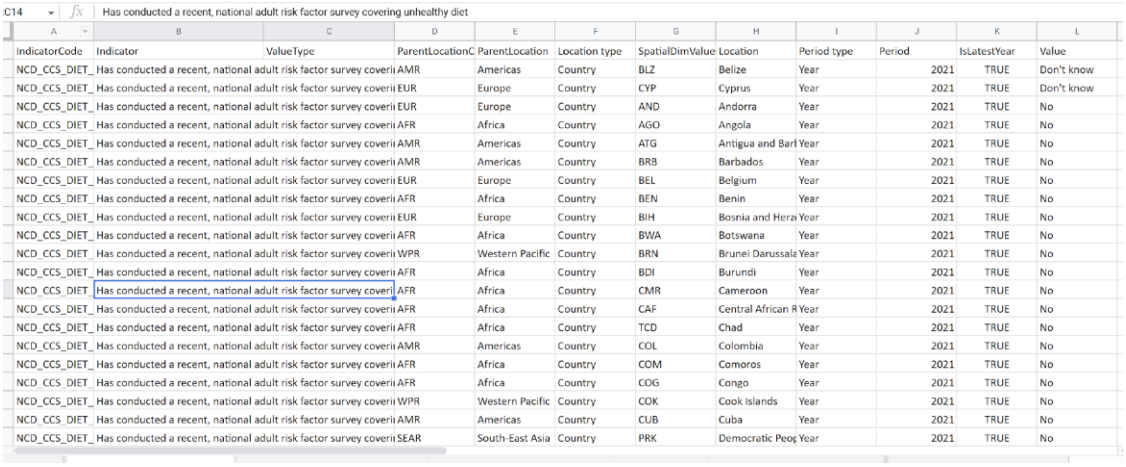
	A	B	C	D	E	F	G	H
		IndicatorCode	ParentLocation	Location	Year	Sex	Age	meanBMI
1		NCD_BMI_MEAN	Africa	Ethiopia	2016	Male	18+ years	20.1
2		NCD_BMI_MEAN	Africa	Eritrea	2016	Male	18+ years	20.2
3		NCD_BMI_MEAN	Africa	Ethiopia	2016	Both sexes	18+ years	20.6
4		NCD_BMI_MEAN	Africa	Eritrea	2016	Both sexes	18+ years	20.7
5		NCD_BMI_MEAN	Africa	Madagascar	2016	Female	18+ years	21.1
6		NCD_BMI_MEAN	Africa	Eritrea	2016	Female	18+ years	21.1
7		NCD_BMI_MEAN	Africa	Ethiopia	2016	Female	18+ years	21.1
8		NCD_BMI_MEAN	Africa	Democratic Repu	2016	Male	18+ years	21.3
9		NCD_BMI_MEAN	Africa	Zambia	2016	Male	18+ years	21.3
10		NCD_BMI_MEAN	Africa	Madagascar	2016	Both sexes	18+ years	21.3
11		NCD_BMI_MEAN	South-East Asia	Timor-Leste	2016	Female	18+ years	21.3
12		NCD_BMI_MEAN	South-East Asia	Timor-Leste	2016	Male	18+ years	21.3
13		NCD_BMI_MEAN	Africa	Rwanda	2016	Male	18+ years	21.3
14		NCD_BMI_MEAN	South-East Asia	Bangladesh	2016	Male	18+ years	21.3
15		NCD_BMI_MEAN	South-East Asia	Timor-Leste	2016	Both sexes	18+ years	21.3
16		NCD_BMI_MEAN	Africa	Niger	2016	Male	18+ years	21.4
17		NCD_BMI_MEAN	Africa	Madagascar	2016	Male	18+ years	21.5
18		NCD_BMI_MEAN	Africa	Burundi	2016	Female	18+ years	21.6
19		NCD_BMI_MEAN	Africa	Uganda	2016	Male	18+ years	21.6
20		NCD_BMI_MEAN	Africa	Burundi	2016	Both sexes	18+ years	21.7
21		NCD_BMI_MEAN	South-East Asia	Bangladesh	2016	Both sexes	18+ years	21.7
22		NCD_BMI_MEAN	South-East Asia	India	2016	Male	18+ years	21.7
23		NCD_BMI_MEAN	Africa	Burundi	2016	Male	18+ years	21.8
24		NCD_BMI_MEAN	Africa	Senegal	2016	Male	18+ years	21.8
25		NCD_BMI_MEAN	Africa	Niger	2016	Both sexes	18+ years	21.8
26		NCD_BMI_MEAN	Western Pacific	Japan	2016	Female	18+ years	21.8
27		NCD_BMI_MEAN	South-East Asia	India	2016	Both sexes	18+ years	21.8
28		NCD_BMI_MEAN	Africa	Chad	2016	Male	18+ years	21.9

Mean BMI After

C2 | Has conducted a recent, national adult risk factor survey covering overweight and obesity

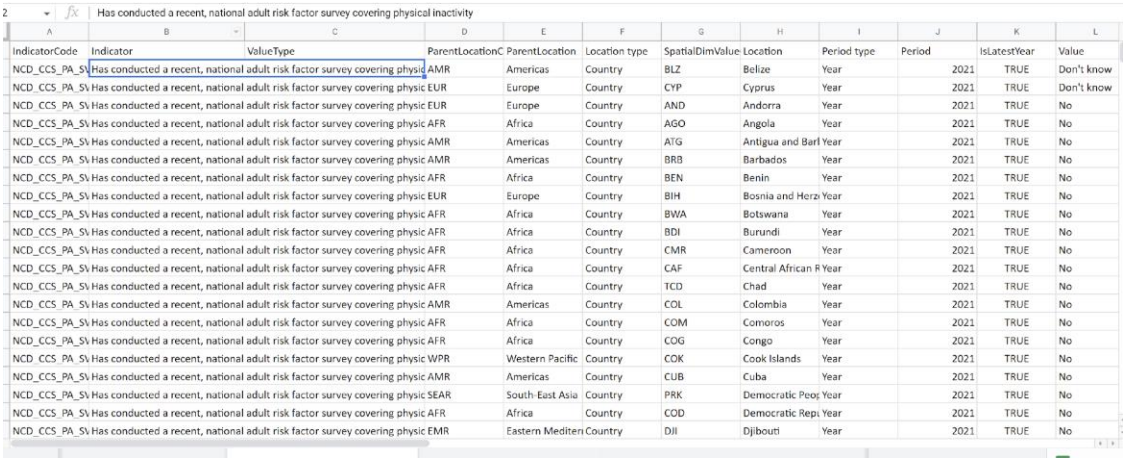
	A	B	C	D	E	F	G	H	I	J	K	L
	IndicatorCode	Indicator	ValueType	ParentLocationCode	ParentLocation	Location type	SpatialDimValue	Location	Period type	Period	IsLatestYear	Value
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		EUR	Europe	Country	CYP	Cyprus	Year	2021	TRUE	Don't know
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		EUR	Europe	Country	AND	Andorra	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	AGO	Angola	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AMR	Americas	Country	ATG	Antigua and Barl	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AMR	Americas	Country	BRB	Barbados	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AMR	Americas	Country	BLZ	Belize	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	BEN	Benin	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		EUR	Europe	Country	BIH	Bosnia and Herz	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	BWA	Botswana	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		WPR	Western Pacific	Country	BRN	Brunei Darussala	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	BDI	Burundi	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	CMR	Cameroon	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	CAF	Central African R	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	TCD	Chad	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AMR	Americas	Country	COL	Colombia	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	COM	Comoros	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	COG	Congo	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		WPR	Western Pacific	Country	COK	Cook Islands	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AMR	Americas	Country	CUB	Cuba	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		SEAR	South-East Asia	Country	PRK	Democratic Peo	Year	2021	TRUE	No
	NCD_CCS_OVWT_SVY	Has conducted a recent, national adult risk factor surve		AFR	Africa	Country	COD	Democratic Rep	Year	2021	TRUE	No

Has conducted a recent, national adult risk factor survey covering overweight and obesity Before



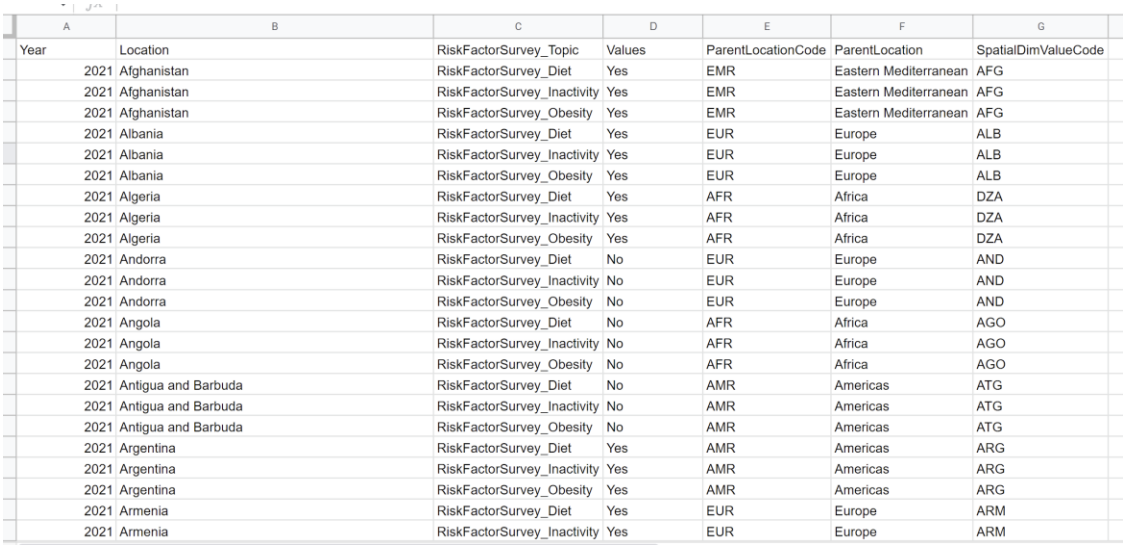
IndicatorCode	Indicator	ValueType	ParentLocationC	ParentLocation	Location type	SpatialDimValue	Location	Period type	Period	IsLatestYear	Value
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AMR	Americas	Country	BLZ	Belize	Year	2021	TRUE	Don't know
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		EUR	Europe	Country	CYP	Cyprus	Year	2021	TRUE	Don't know
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		EUR	Europe	Country	AND	Andorra	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	AGO	Angola	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AMR	Americas	Country	ATG	Antigua and Barl	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AMR	Americas	Country	BRB	Barbados	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		EUR	Europe	Country	BEL	Belgium	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	BEN	Benin	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		EUR	Europe	Country	BIH	Bosnia and Herz	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	BWA	Botswana	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		WPR	Western Pacific	Country	BRN	Brunei Darussala	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	BDI	Burundi	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	CMR	Cameroon	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	CAF	Central African R	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	TCD	Chad	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AMR	Americas	Country	COL	Colombia	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	COM	Comoros	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AFR	Africa	Country	COG	Congo	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		WPR	Western Pacific	Country	COK	Cook Islands	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		AMR	Americas	Country	CUB	Cuba	Year	2021	TRUE	No
NCD_CCS_DIET_	Has conducted a recent, national adult risk factor survey coveri		SEAR	South-East Asia	Country	PRK	Democratic Peo	Year	2021	TRUE	No

Has conducted a recent, national adult risk factor survey covering unhealthy diet Before



IndicatorCode	Indicator	ValueType	ParentLocationC	ParentLocation	Location type	SpatialDimValue	Location	Period type	Period	IsLatestYear	Value
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AMR	Americas	Country	BLZ	Belize	Year	2021	TRUE	Don't know
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		EUR	Europe	Country	CYP	Cyprus	Year	2021	TRUE	Don't know
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		EUR	Europe	Country	AND	Andorra	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	AGO	Angola	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AMR	Americas	Country	ATG	Antigua and Barl	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AMR	Americas	Country	BRB	Barbados	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	BEN	Benin	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		EUR	Europe	Country	BIH	Bosnia and Herz	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	BWA	Botswana	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	BDI	Burundi	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	CMR	Cameroon	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	CAF	Central African R	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	TCD	Chad	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AMR	Americas	Country	COL	Colombia	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	COM	Comoros	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	COG	Congo	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		WPR	Western Pacific	Country	COK	Cook Islands	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AMR	Americas	Country	CUB	Cuba	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		SEAR	South-East Asia	Country	PRK	Democratic Peo	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		AFR	Africa	Country	COD	Democratic Rep	Year	2021	TRUE	No
NCD_CCS_PA_SV	Has conducted a recent, national adult risk factor survey covering physic		EMR	Eastern Mediter	Country	DJI	Djibouti	Year	2021	TRUE	No

Has conducted a recent, national adult risk factor survey covering physical inactivity Before

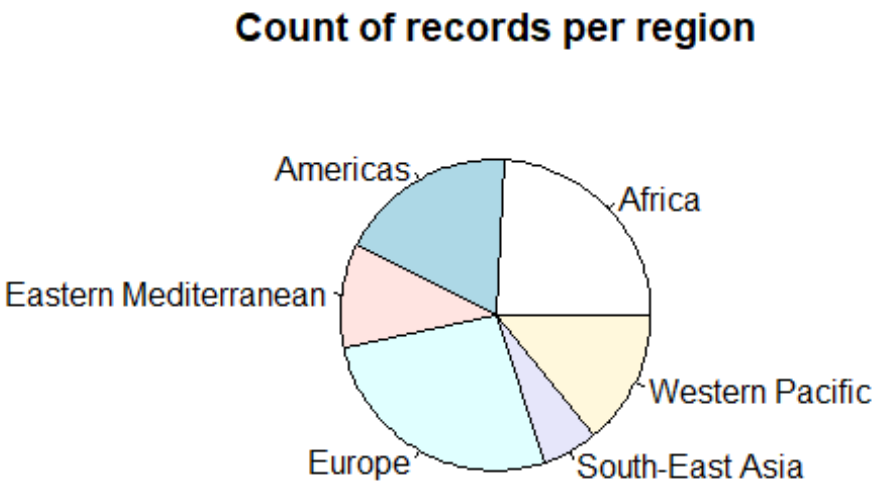


Year	Location	RiskFactorSurvey_Topic	Values	ParentLocationCode	ParentLocation	SpatialDimValueCode
2021	Afghanistan	RiskFactorSurvey_Diet	Yes	EMR	Eastern Mediterranean	AFG
2021	Afghanistan	RiskFactorSurvey_Inactivity	Yes	EMR	Eastern Mediterranean	AFG
2021	Afghanistan	RiskFactorSurvey_Obesity	Yes	EMR	Eastern Mediterranean	AFG
2021	Albania	RiskFactorSurvey_Diet	Yes	EUR	Europe	ALB
2021	Albania	RiskFactorSurvey_Inactivity	Yes	EUR	Europe	ALB
2021	Albania	RiskFactorSurvey_Obesity	Yes	EUR	Europe	ALB
2021	Algeria	RiskFactorSurvey_Diet	Yes	AFR	Africa	DZA
2021	Algeria	RiskFactorSurvey_Inactivity	Yes	AFR	Africa	DZA
2021	Algeria	RiskFactorSurvey_Obesity	Yes	AFR	Africa	DZA
2021	Andorra	RiskFactorSurvey_Diet	No	EUR	Europe	AND
2021	Andorra	RiskFactorSurvey_Inactivity	No	EUR	Europe	AND
2021	Andorra	RiskFactorSurvey_Obesity	No	EUR	Europe	AND
2021	Angola	RiskFactorSurvey_Diet	No	AFR	Africa	AGO
2021	Angola	RiskFactorSurvey_Inactivity	No	AFR	Africa	AGO
2021	Angola	RiskFactorSurvey_Obesity	No	AFR	Africa	AGO
2021	Antigua and Barbuda	RiskFactorSurvey_Diet	No	AMR	Americas	ATG
2021	Antigua and Barbuda	RiskFactorSurvey_Inactivity	No	AMR	Americas	ATG
2021	Antigua and Barbuda	RiskFactorSurvey_Obesity	No	AMR	Americas	ATG
2021	Argentina	RiskFactorSurvey_Diet	Yes	AMR	Americas	ARG
2021	Argentina	RiskFactorSurvey_Inactivity	Yes	AMR	Americas	ARG
2021	Argentina	RiskFactorSurvey_Obesity	Yes	AMR	Americas	ARG
2021	Armenia	RiskFactorSurvey_Diet	Yes	EUR	Europe	ARM
2021	Armenia	RiskFactorSurvey_Inactivity	Yes	EUR	Europe	ARM

Global Surveys of Overweight/Obese, Unhealthy Diet, and Physical Inactivity After

Data Exploration

In our exploratory analysis we are beginning to ask questions about the global status of overweight and obesity. Using data from the World Health Organization we are able to dive into the rates of obesity across different regions of the world.




```
## integer(0)
```

Figure 1: The World Health Organization's data on Prevalence of Overweight and Obesity in the regions is not sampled uniformly across the regions. Some regions had more countries participating in the research compared to others.

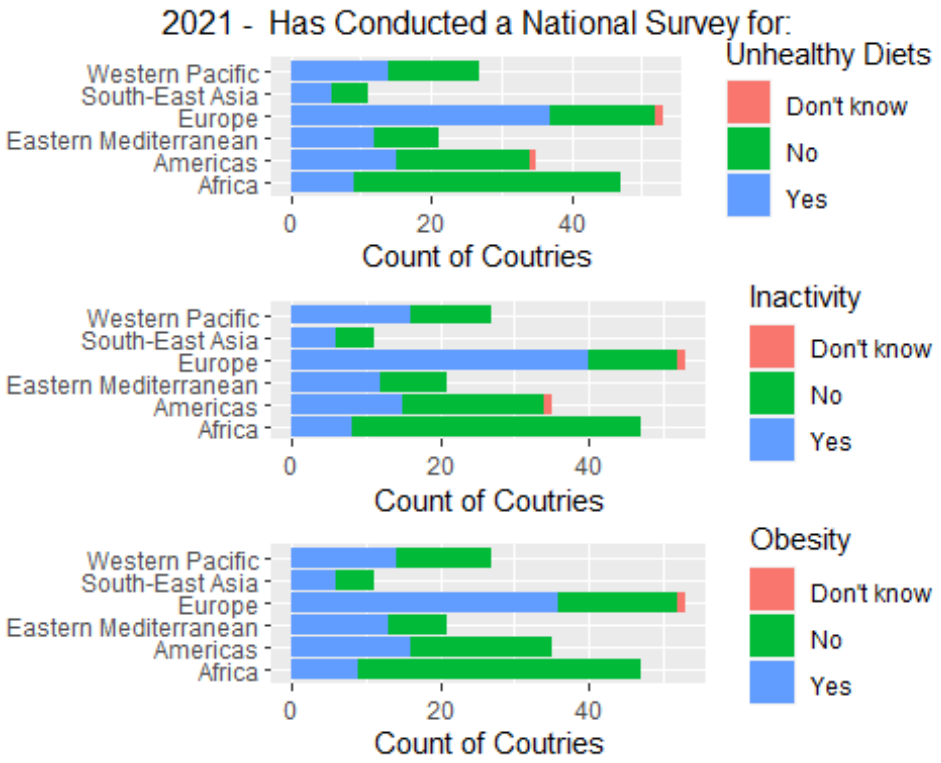


Figure 2: The above graph shows the count of countries in each of the regions that are taking surveys or not based on the three topics. These three topics are:

- Has conducted a recent, national adult risk factor survey covering overweight and obesity
- Has conducted a recent, national adult risk factor survey covering unhealthy diet
- Has conducted a recent, national adult risk factor survey covering physical inactivity

This graph does not necessarily show which regions are the most active or have the least healthy of diets. It is more an understanding of which regions are trying to understand the population's trends.

The World Health Organization has many years of data on prevalence of overweight among adults and prevalence of obesity among adults. The measurement method for deterimining "overweight" and "obese" is using the Body Mass Index. According to the United States CDC, "Body Mass Index (BMI) is a person's weight in kilograms (or pounds) divided by the square of height in meters (or feet). A high BMI can indicate high body fatness. BMI screens for weight categories that may lead to health problems, but it does not diagnose the body fatness or health of an individual". Some questions to ask are:

1. Do overweight and obesity rates depend on the region?
2. Is there a difference in the prevalence of overweight and obesity among different sexes?
3. How has the prevalence of overweight and obesity changed over the years?

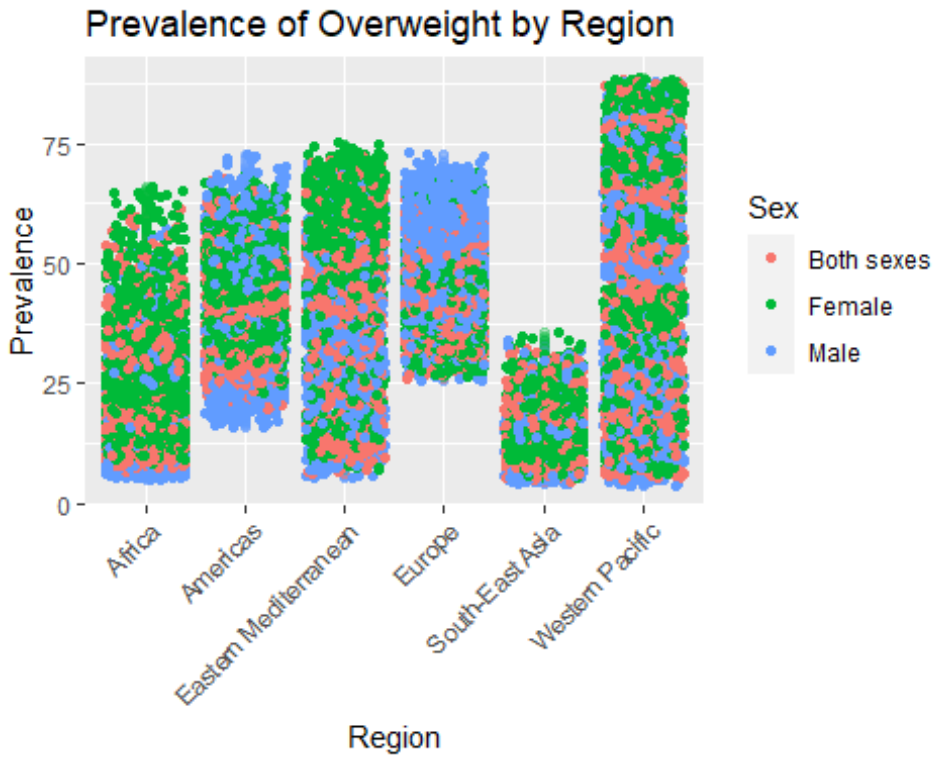


Figure 3

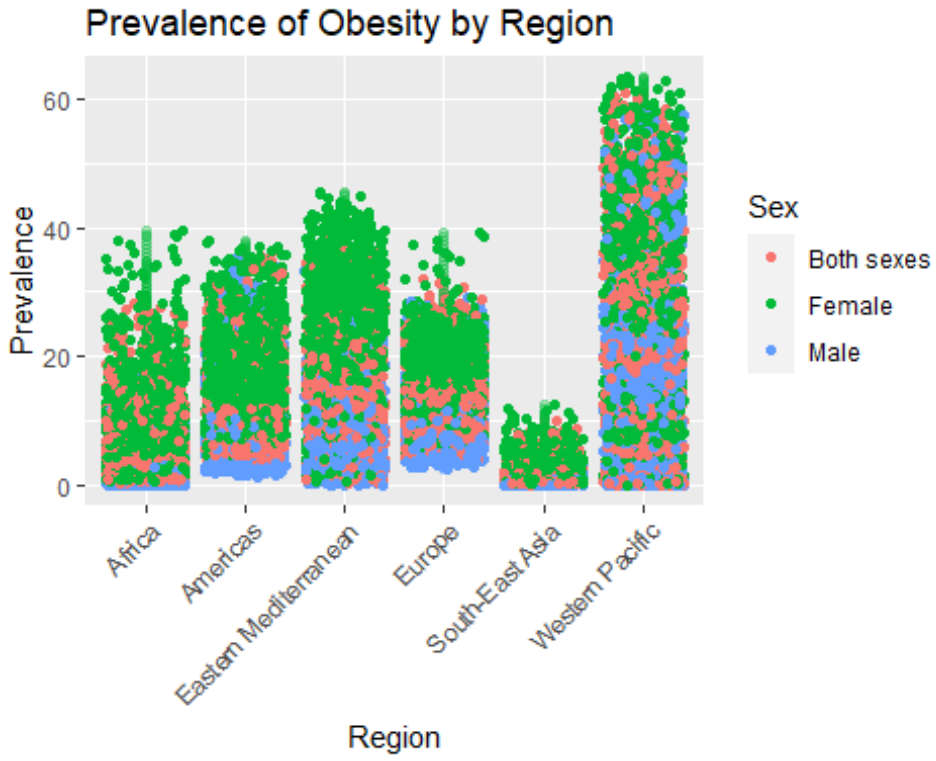


Figure 4

As we can see in Figure 3 and Figure 4, the prevalence of overweight and obesity might have different distributions based on region. We can also see that the overall maximum percentages are lower for obesity than for overweight. This is a good thing, and it is also to be expected. In order to cross the boundary into obese one must first be overweight. Figure 3 also seems to have a more even spread of values for the different sexes, however in Figure 4 females seem to have a higher prevalence of obesity compared to males of the same region.

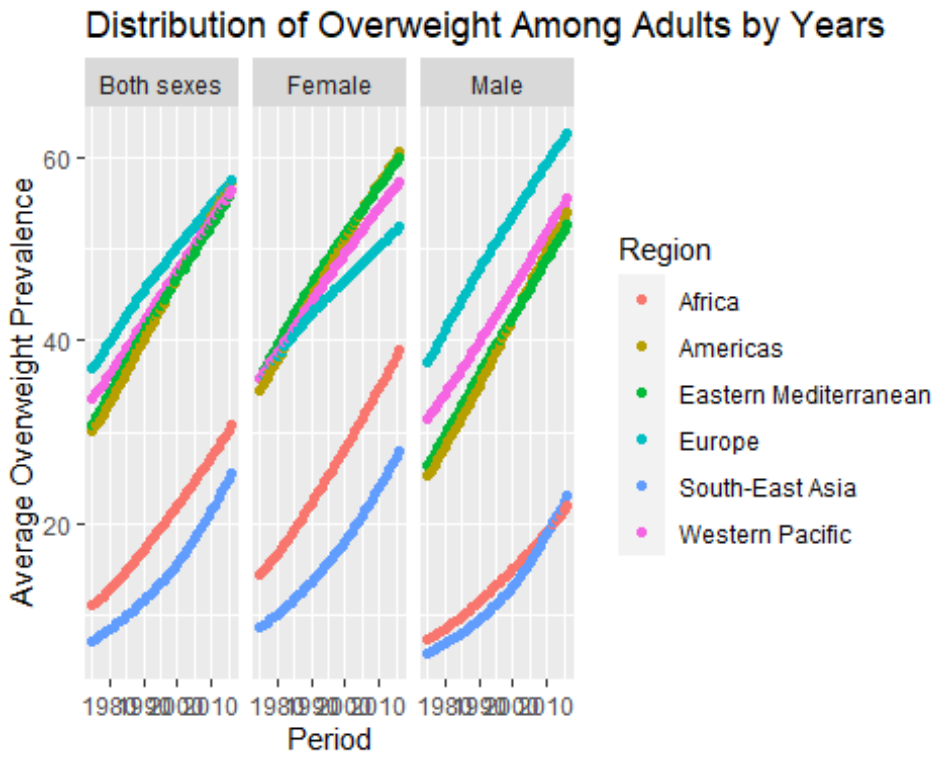


Figure 5

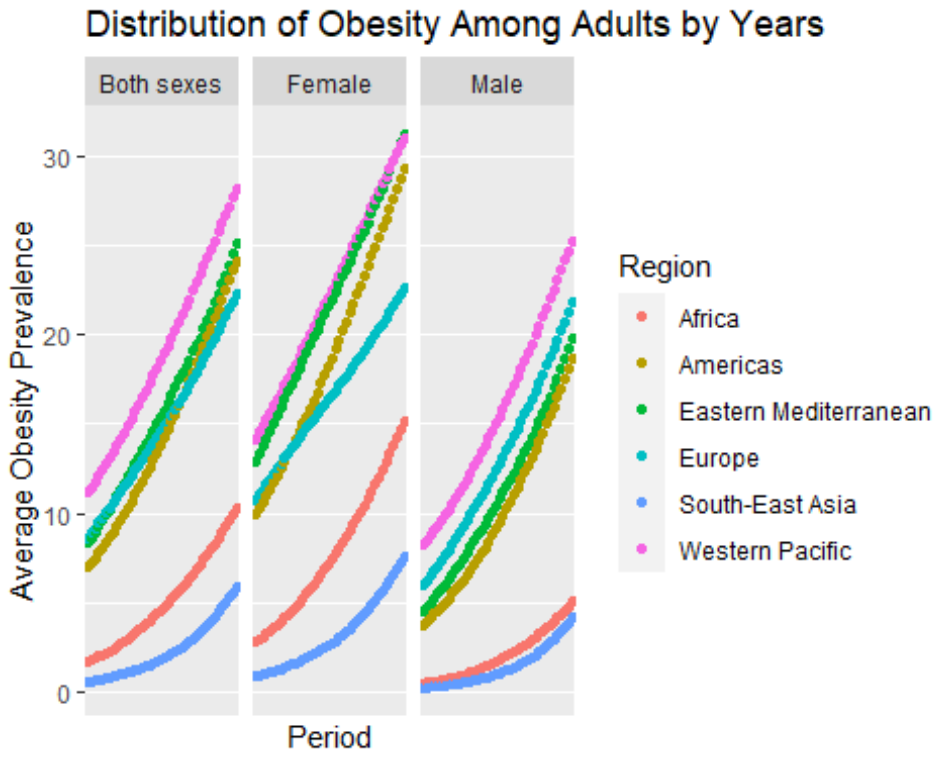


Figure 6

One of our questions was the distribution of the prevalences over the years. From our exploratory analysis we can see that for both genders (male and female), the prevalence of overweight and obesity in all of the regions participating in the study showed an increase over the time period 1975 to 2016. Some regions showed a quicker increase than others. This can prompt further research into asking why these regions have such a high percent increase in obesity and overweight rates.

```
##    Statistic          BMI  Overweight  Obesity Blood Pressure Depression
## 1       Mean        25.78       48.44    19.51          24.43       4.43
## 2     Median         26.3        54.2     20.2           24.5        4.4
## 3       Mode         26.5        53.9     19.3           19.7       5.07
## 4      Range 20.5 - 32.9 17.7 - 88.3 2 - 60.7      11 - 33.4 2.9 - 6.31
## 5        Var         5.66      291.17   127.98           20.4       0.45
## 6    Std Dev         2.38       17.06    11.31           4.52       0.67
## 7     25th Q         23.9       30.65     9.15          21.15       3.96
## 8     50th Q         26.3        54.2     20.2           24.5        4.4
## 9     75th Q         27.1       60.15     25.2          28.25      4.955
## 10       IQR          3.2        29.5    16.05            7.1      0.995
##          Deaths
## 1         20.71
## 2         21.05
## 3         15.57
## 4  8.4 - 51.73
## 5         63.81
```

```
## 6          7.99
## 7         15.18
## 8         21.05
## 9         24.77
## 10         9.59
```

Figure 7: Statistics Table for Quantitative Data A concise table for quantitative data's measures of center.

Distribution Analysis

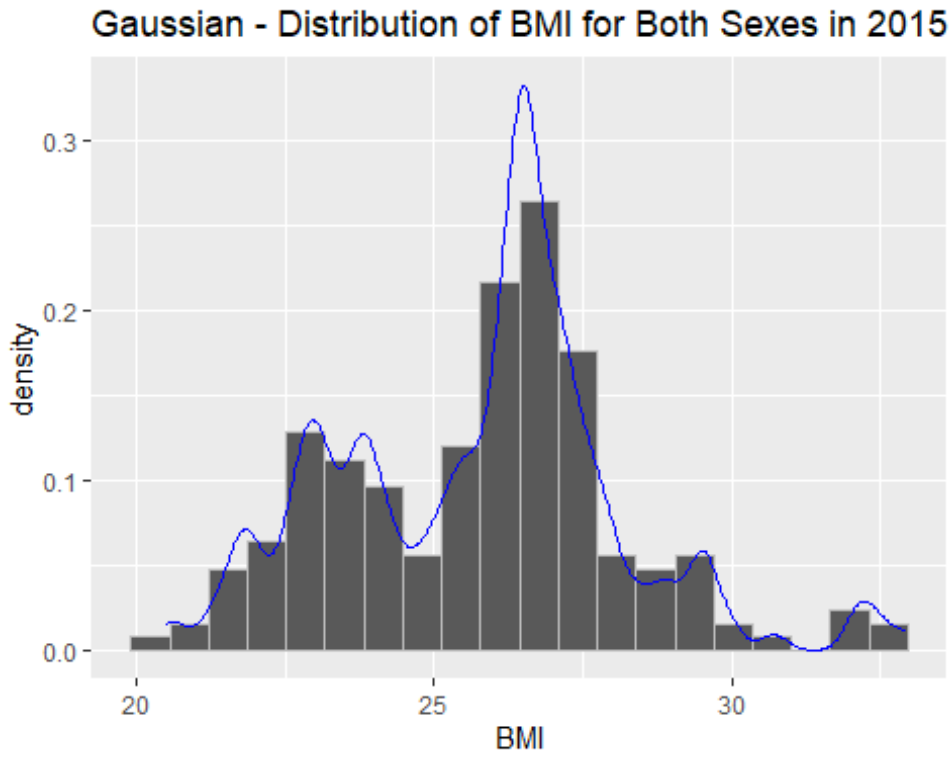


Figure 8

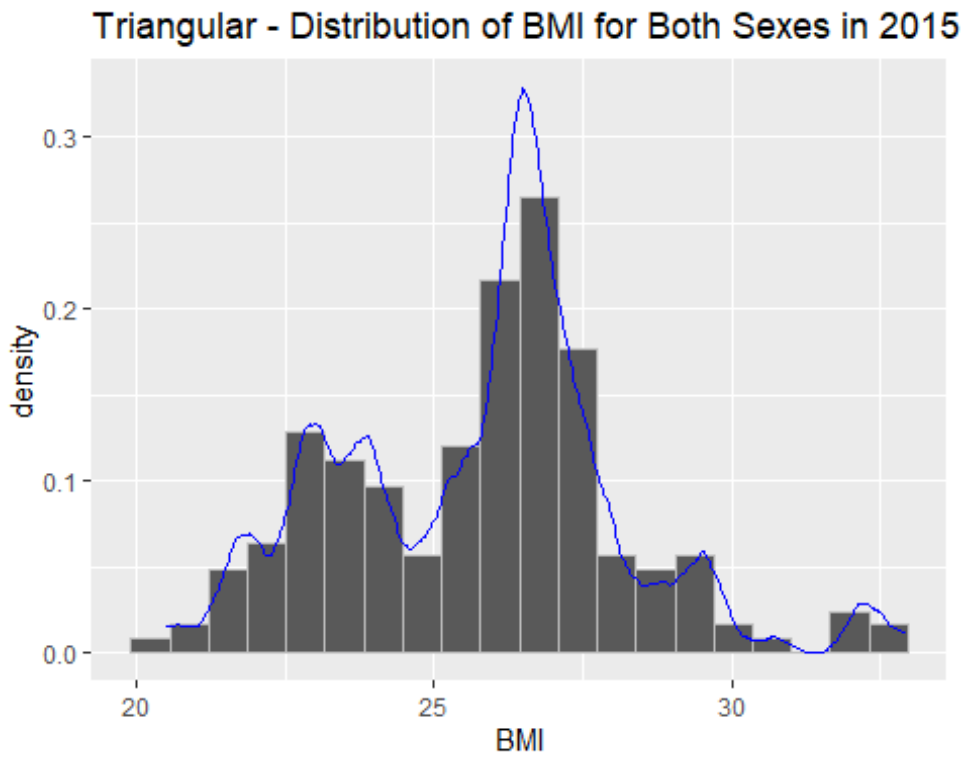


Figure 9

As we can see in figure 8 and figure 9, we used the kernel estimators. We found that the BMI for both sexes in 2015 is bimodal. In this case, we used Gaussian and Triangular kernel functions for the density plots. It peaks at approximately 26. These visualizations show that these estimators look practically the same.

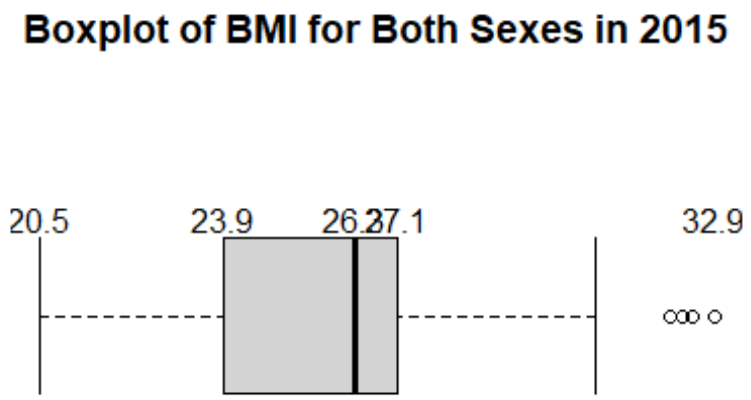


Figure 10: From the boxplot, there are a few outliers on the right. The boxplot shows that the median is 26.3. The quartile 1 is 23.9 and the quartile 3 is 27.1.

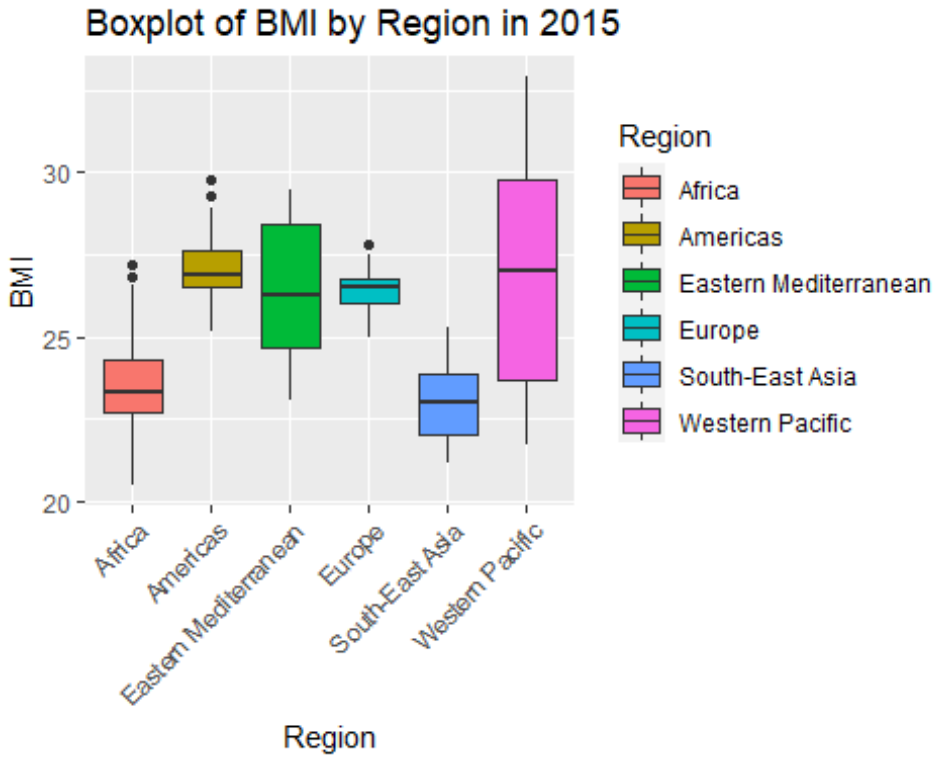


Figure 11: The above graph shows that America and Western Pacific have the highest median of BMI by region which are followed by Europe, Eastern Mediterranean, and Africa respectively. South-East Asia has the lowest median of BMI among these region.

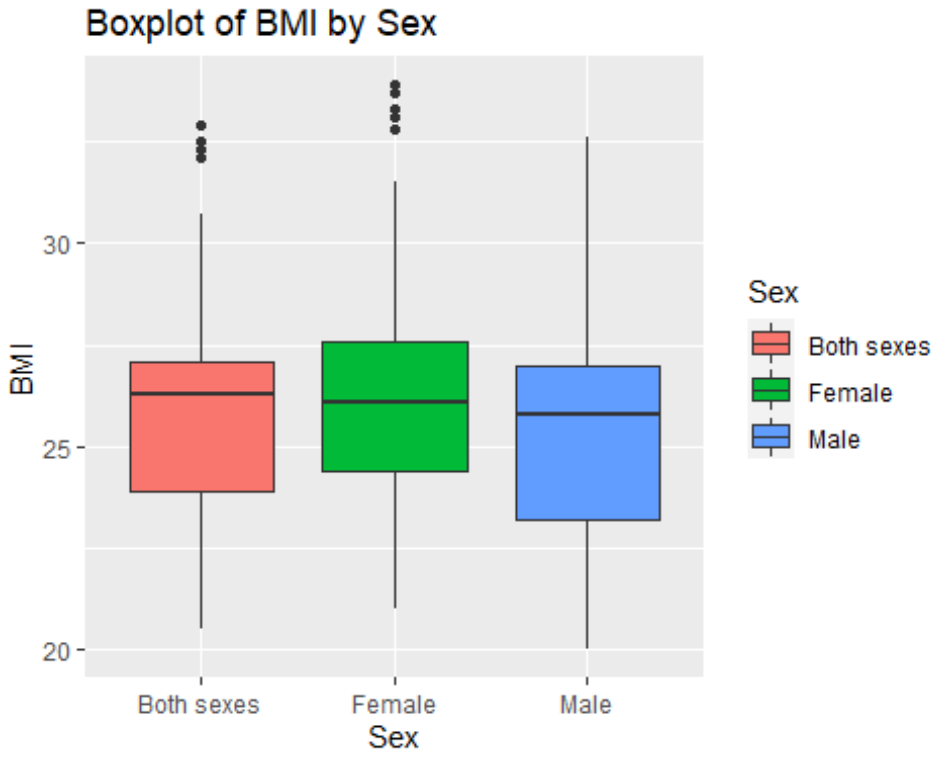


Figure 12: The boxplot shows that female has slightly higher median BMI than male. The median BMI of female, male, and both sexes are 26.1, 25.8, and 26.3 respectively.

```
# Histogram Overweight Both Sexes 2015
ggplot(data = data_2015, aes(x=Prev_overwt)) +
  geom_histogram(aes(y=..density..), bins = 35) +
  geom_density(color= 'red') + xlim(0,65)  + xlab("Prevalence of Overweight")
+
  ggtitle("Distribution of Overweight World-Wide in 2015")
```

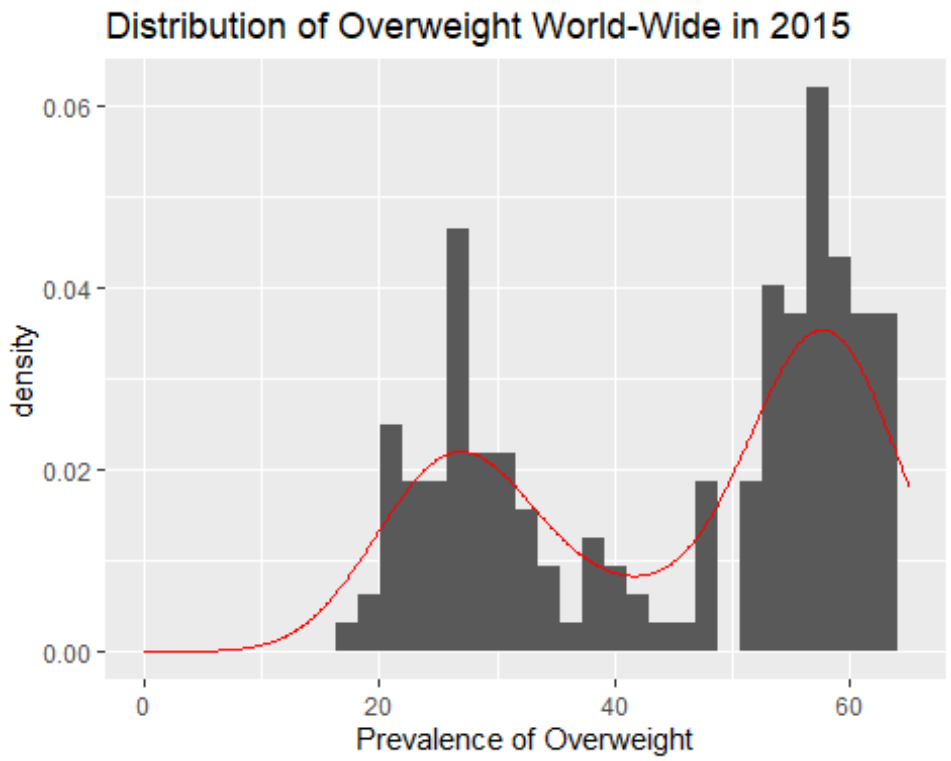


Figure 13: The prevalence of overweight consists of the percentage of individuals within the population of a certain country that have BMI greater than 25. This means that obese are individuals are counted in with the overweight population. This is something important to keep in mind as we look at distribution of the prevalence of overweight people. From figure 13 we see that the distribution is bimodal with peaks at around 30 and 60 percent, the latter of which being the largest peak. There is quite a significant dip between these two peaks suggesting that in large countries are either fairly healthy or fairly overweight and not much in between.

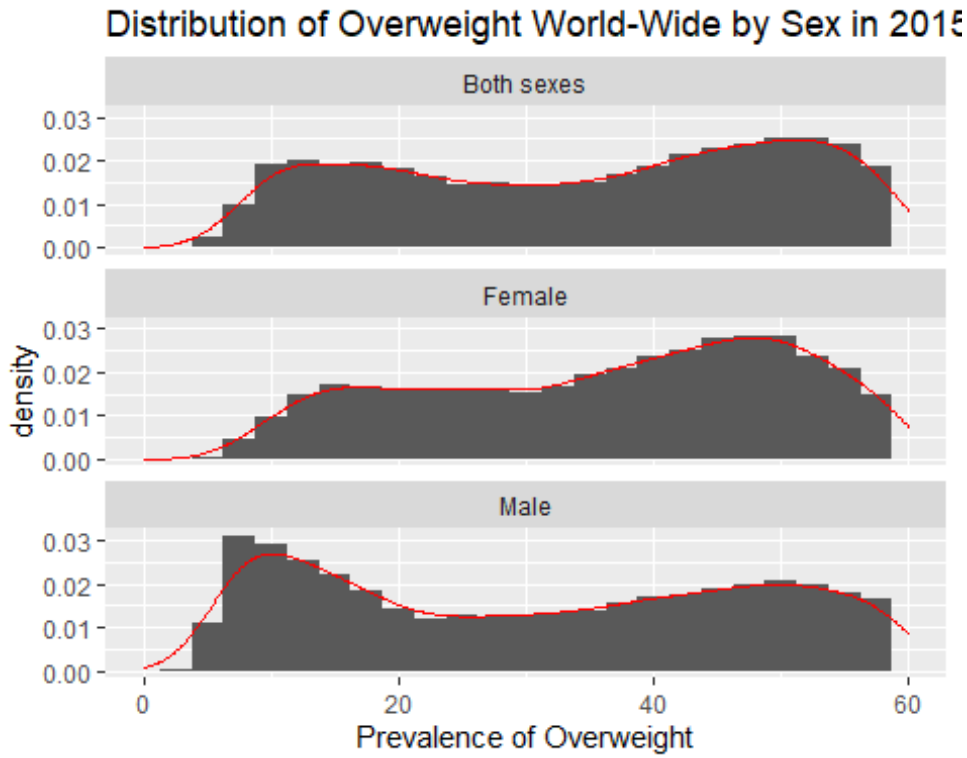


Figure 14: The prevalence of overweight individuals by sex is largely very similar. Both sexes have bimodal distribution however there are slight differences. There are more countries wear males have lower prevalence of obesity shown by a strong peak at around 20 percent. Whereas the strongest peak for the female distribution is closer to 50 percent. This suggests that in general there is more prevalence of overweight females than males.

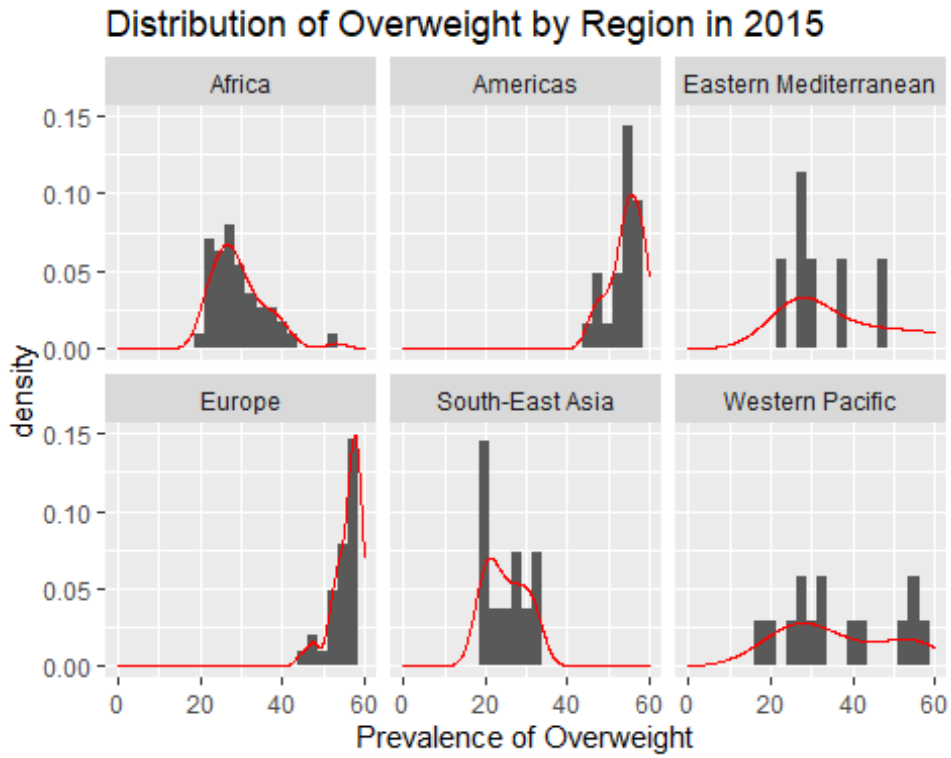


Figure 15: Prevalence of overweight individuals varies significantly by region. The Americas and Europe have the highest rates of overweight amongst the six regions. In the Eastern Mediterranean and Western pacific regions we see a more uniform distribution of overweight countries. Countries in Africa and South-East Asia have the smallest prevalences of overweight people with values ranging from 20-40 percent.

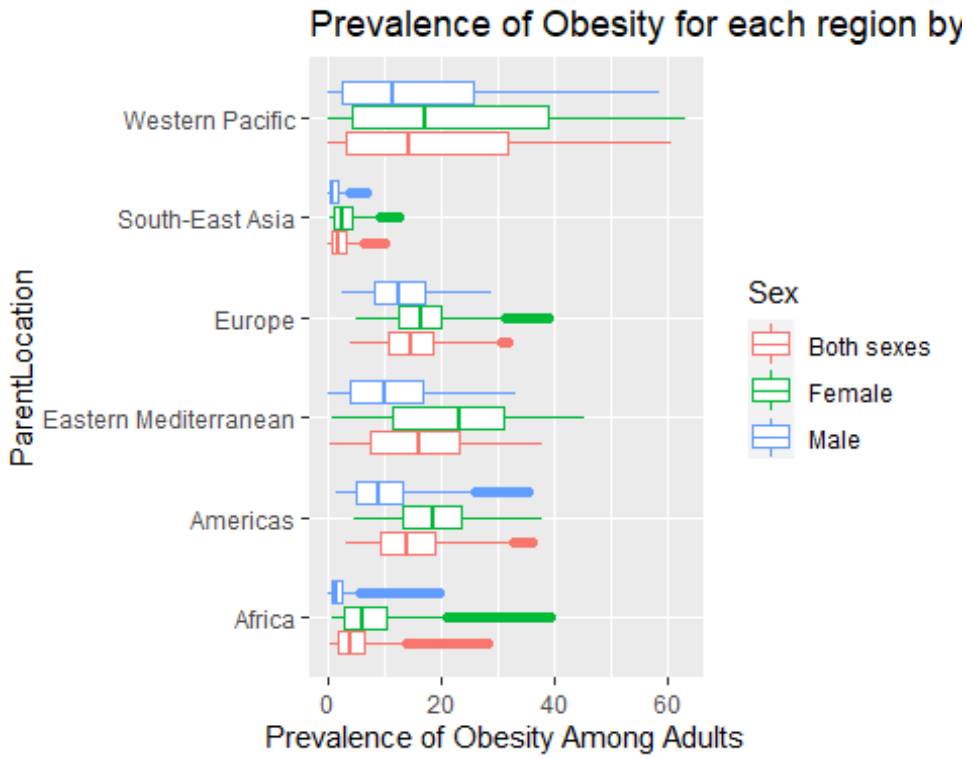


Figure 16: In figure 16 we can see the shapes of our distributions by both region and sex. This multi faceted box-plot shows the variation in the prevalence of obesity based on these two factors. In the following plots we will analyze each one of these characteristics individually to uncover the relationships between sex, region and obesity.

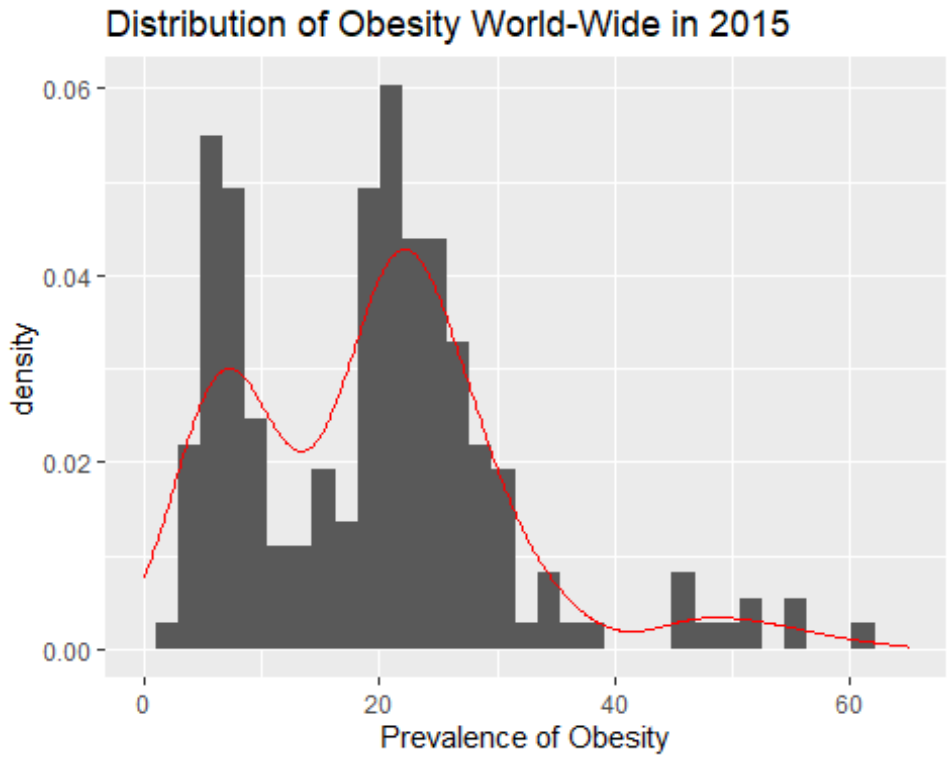


Figure 17: In the figure above we can see the distribution of the prevalence of obesity in countries around the world in the year 2015. This plot accounts for both males and females. The density estimate line shows us that the distribution is bimodal with peaks at 7 and 20 percent obesity rate. In addition to this we see a heavy positive skew that shows that there are countries that have rates of obesity that exceed 50 percent of the population.

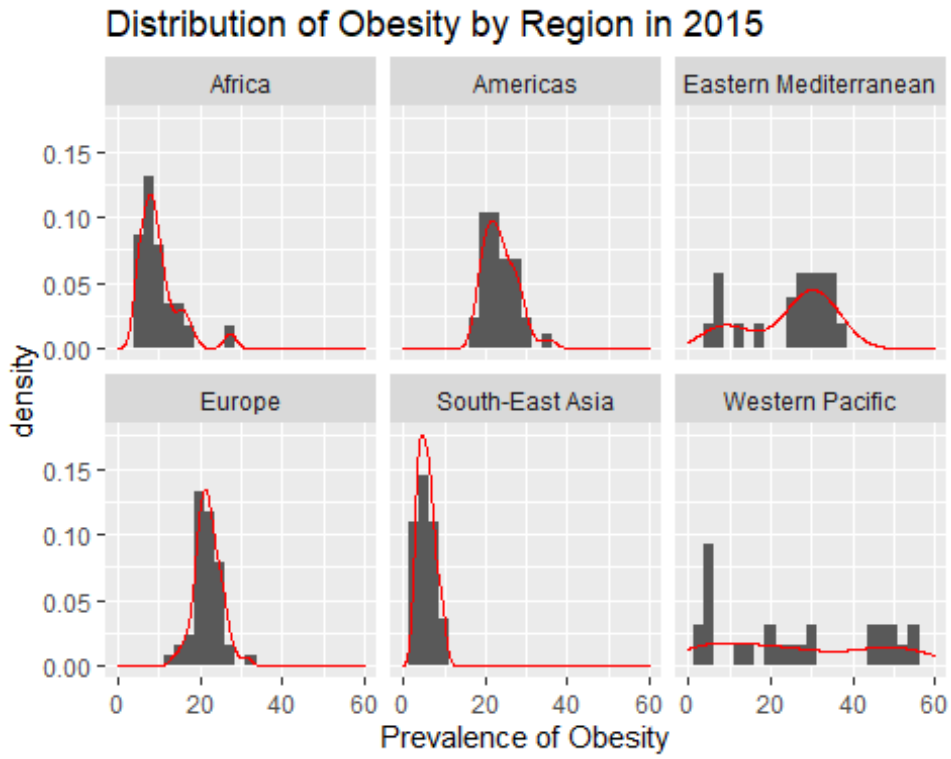


Figure 18: In figure 18 we have the distribution of obesity by region in the world accounting for both males and females. This chart shows that there are distinct differences in the prevalence of obesity between regions. Africa and South-East Asia have many more countries with lower rates of obesity having peaks in their distributions at 10 and 5 percent respectively. In contrast to this the Americas, Eastern Mediterranean, European, and Western Pacific regions have distributions that are centered much further to the right showing that the countries in these regions tend to have much higher rates of obesity.

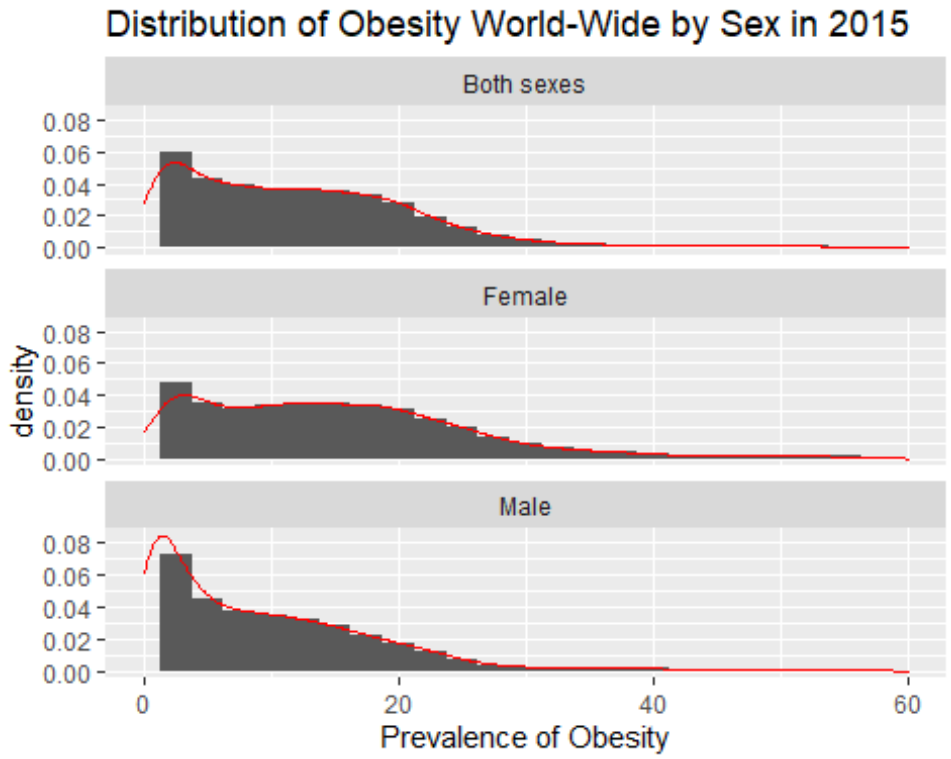


Figure 19 In figure 19 we see another breakdown of the distribution of obesity but this time by sex. Both males and females have a bimodal distribution of obesity. However, the female data is slightly more spread out showing that there are more countries that have a male prevalence of obesity over 20 percent. The male distribution has a fairly heavy peak between 5-10 percent obesity. This shows us that males have slightly lower rates of obesity than females.

Overall, we have observed patterns in the distributions of the prevalence of obesity that suggest that there are many factors that play a role in obesity rates around the world. Both sex and region seem to be key in analyzing how obesity affects countries world-wide.

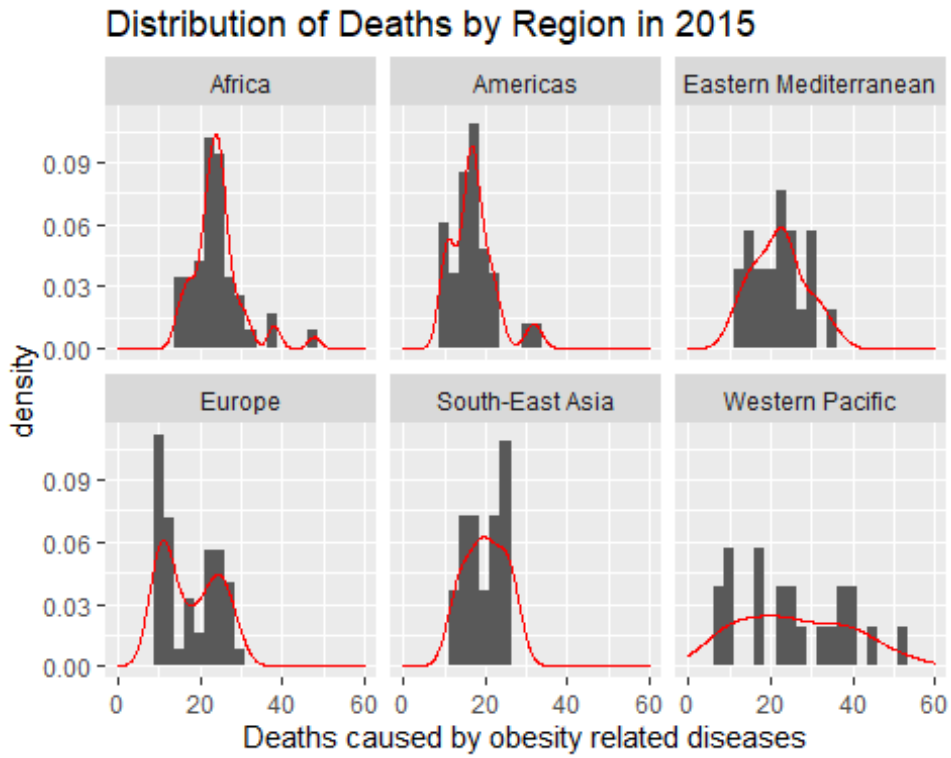
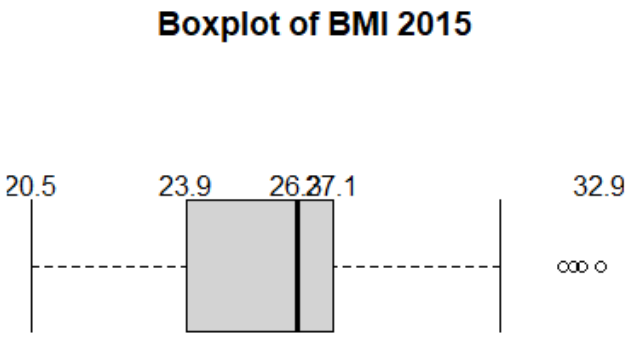
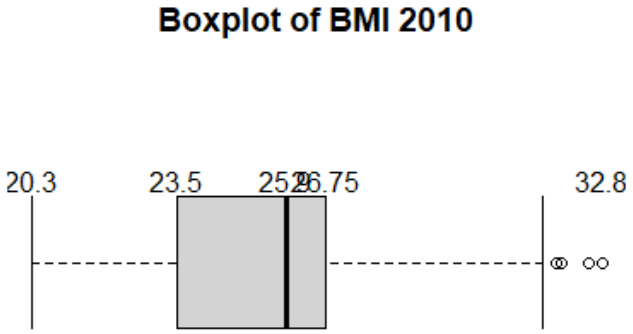


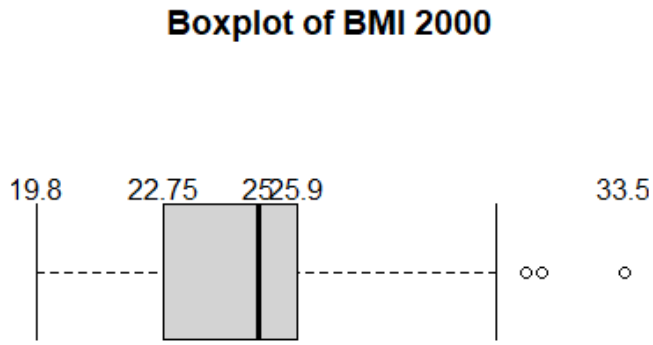
Figure 20 In figure 20, we can see that in Africa, Americas, Eastern Mediterranean, and South-East Asia the distribution of deaths follows a uni-modal distribution. Also Europe and Western Pacific have a bimodal distribution for their region's probability of deaths. Later we wish to analyze if the means of the different regions differ significantly or not through a hypothesis test.

Results

Variable Case Study





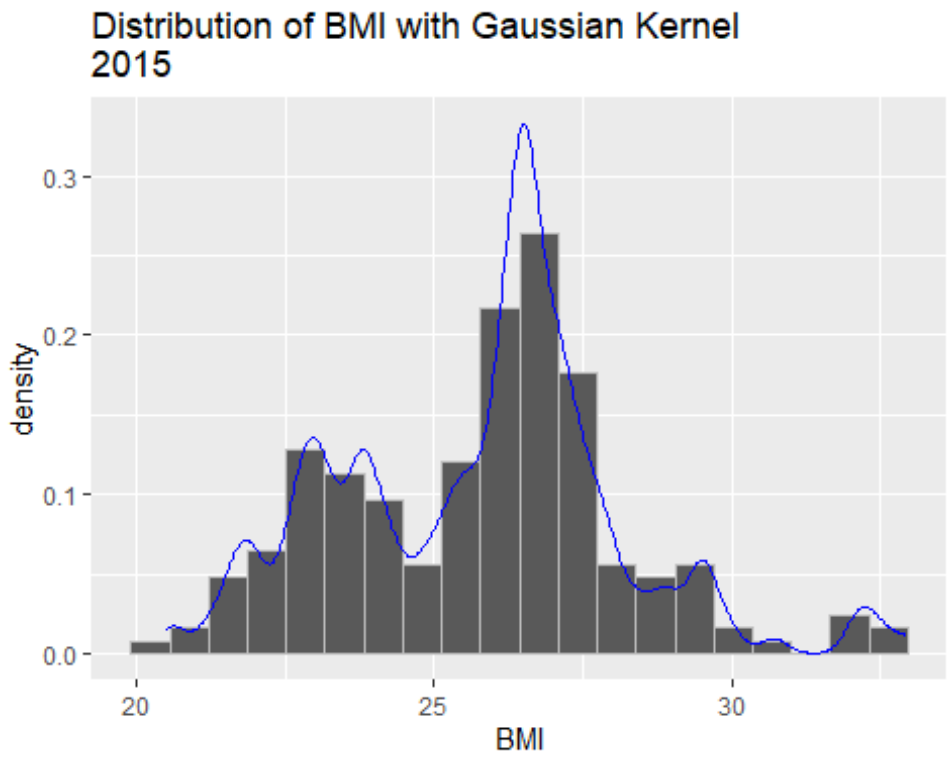


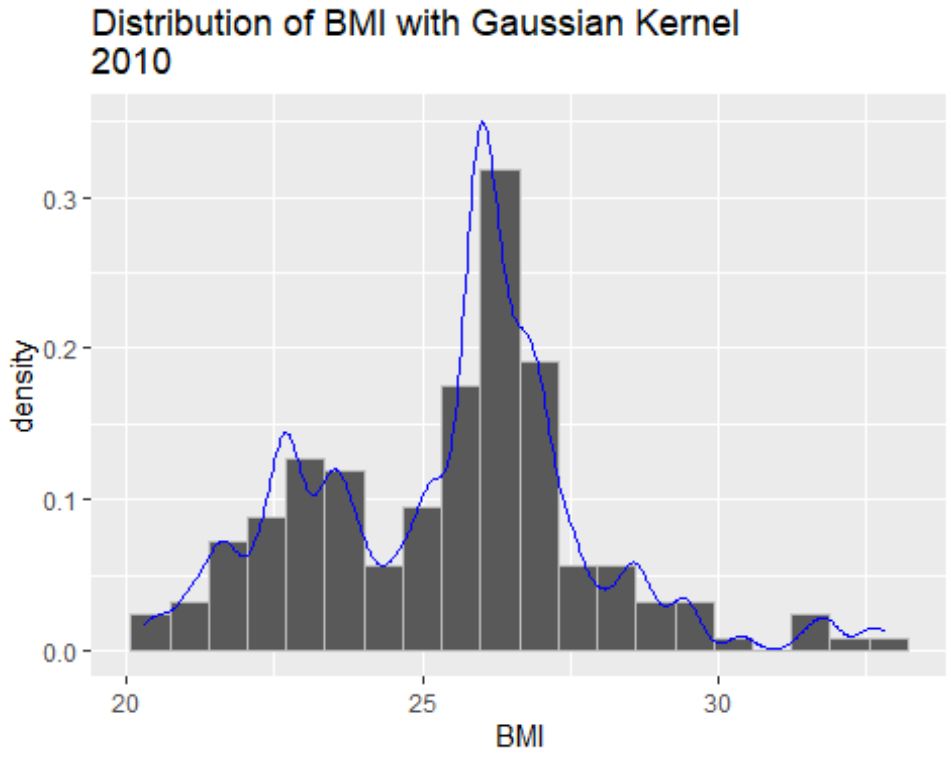
Throughout this paper we discussed the population distribution for prevalence of specific BMI's. For the "variable case study" we wish to analyze the BMI value across the different countries.

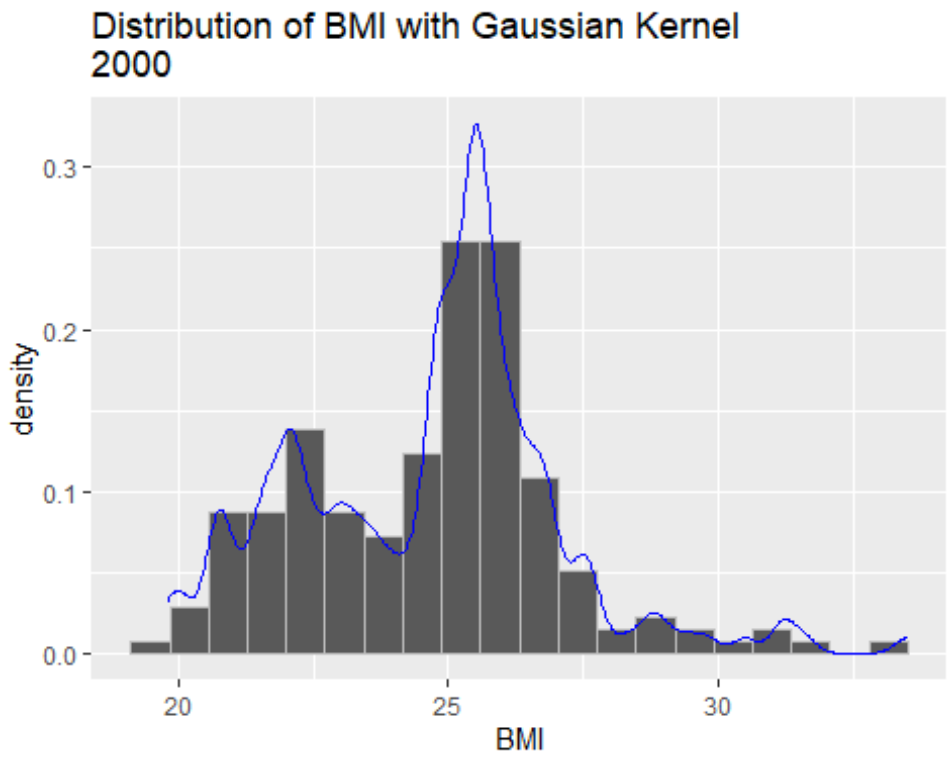
BMI is a continuous value from 0 to infinity. However infinity isn't reasonable in real applications as there are no individuals with infinite weight or height. 0 is also an "unreasonable" metric because this would require the individual to have 0 mass (i.e. they don't exist). Since we are not dealing with a population that has ghosts and giants to sample from these are unreasonable numbers. Though based on the equation to find BMI, they're not mathematically impossible.

Earlier in our distribution analysis, we first talked about the BMI distribution for all the countries in the study in 2015. We would say we sampled the entire population, however, this is untrue because not all countries in the world participated in the World Health Organization's study.

The box plots for the countries average BMI for years 2015, 2010, and 2000 show that over the years the average BMI of all the countries in the study were increasing.

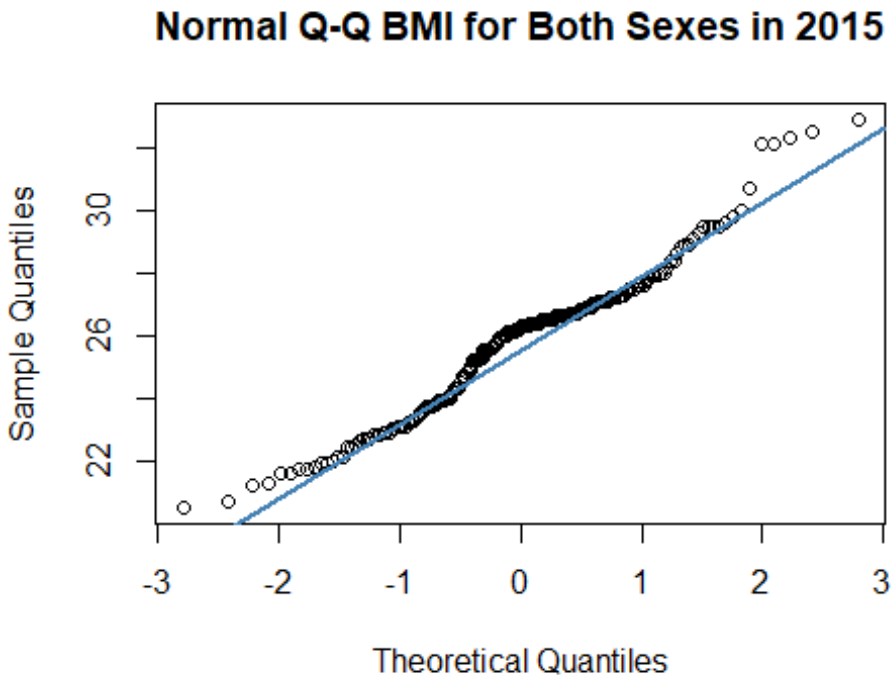

Distribution of BMI with Gaussian Kernel
2015
density
0.0
0.1
0.2
0.3
20
25
30
BMI


Distribution of BMI with Gaussian Kernel
2010
density
0.0
0.1
0.2
0.3
20
25
30
BMI



We can also see from the histograms to the above and to the left that the distribution of the mean BMI's stayed relatively similar between the years 2015, 2010, and 2000. This distribution isn't following any of the standard distributions that are commonly known. That is because the distribution is bimodal with two consistent peaks. One major peak around the mean and median and then a smaller peak on the lower side of the distribution.

Also seen from the box plots in the histograms we can see that there are outliers in the upper ranges of the distribution, however, this is not as much the case in the lower. As we can see from the Statistical Table in Figure 7, the mean is pulled higher than the median.



The Normal QQ-Plot to the left shows that the distribution of BMI averages for the countries does not follow a normal distribution. However, a normal distribution could be used as a rough estimate of the distribution, if needed. As the values go past 2 standard deviations away from the mean the estimation would be less accurate.

Hypothesis Testing

Two Sample One Sided T test

Ho: Mean of Prevalence of Obesity in 2015 = Mean of Prevalence of Obesity in 2014/2013/2012/2011/2010

Ha: Mean of Prevalence of Obesity in 2015 > Mean of Prevalence of Obesity in 2014/2013/2012/2011/2010

Results from T test

Years	2015 vs 2014	2015 vs 2013	2015 vs 2012	2015 vs 2011	2015 vs 2010
p-value	0.2597194	0.102228	0.0293684	0.0061888	0.0009315

```
##
##  Welch Two Sample t-test
##
## data:  data2015$Prev_overwt and data2012$Prev_overwt
```

```
## t = 1.6101, df = 1144, p-value = 0.05383
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.03806192         Inf
## sample estimates:
## mean of x mean of y
##  48.35899  46.66091
```

In order to determine whether or not obesity is becoming more prevalent over time we conducted multiple two sample one-sided t tests comparing 2015 to the 5 previous years. We are testing the null hypothesis that the mean prevalence of obesity in 2015 is equal to the mean prevalence of obesity in the years 2010-2014. The alternative hypothesis is that the mean prevalence of obesity in 2015 is greater than the mean in the years 2010-2014. In the table above we see the results from our 5 tests. As shown in table above, when comparing the years 2015 and 2012 we see our first statistically significant result using alpha=0.05. With a p-value of 0.029 we reject the null hypothesis and conclude that the prevalence of obesity in 2015 is greater than the prevalence of obesity in 2012. This reaffirms our exploratory analysis and shows that obesity is increasing worldwide as time moves on.

ANOVA test

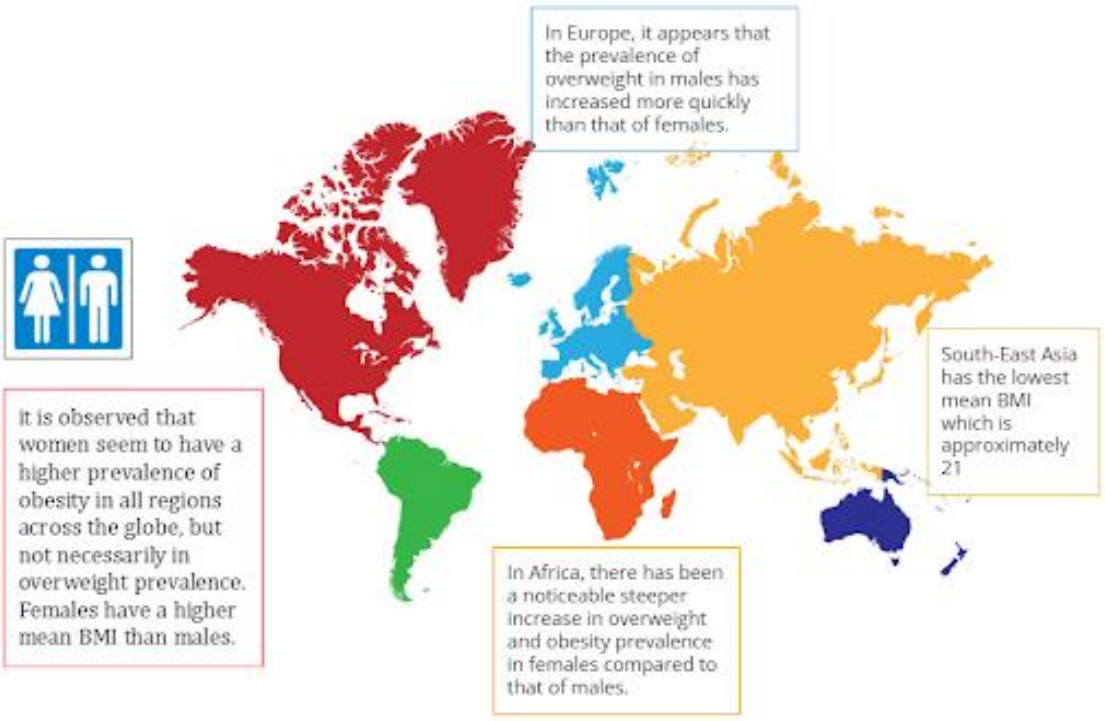
Ho: Mean of Deaths in America= Mean of Deaths in Africa = Mean of Deaths in Southeast Asia = Mean of Deaths in Europe = Mean of Deaths in Western Pacific = Mean of Deaths in Eastern Mediterranean

Ha: Not all are equal

```
##                 Df Sum Sq Mean Sq F value   Pr(>F)
## ParentLocation   5   2164   432.8   8.107 6.56e-07 ***
## Residuals      177   9449    53.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Moving forward with our analysis, we conducted an anova test with the goal of determining whether or not the probability that an individual will die between age 30 and 70 from any cardiovascular disease, cancer, diabetes or chronic respiratory disease differs by region of the world. With a p-value 6.56e-7 and alpha=0.05 we reject the null hypothesis and conclude that at least one region differs from the others. This means that region has a significant impact on the likelihood of death due to obesity related diseases.

Conclusion



In the end, we can conclude that there is a steady increase in obesity and overweight across the globe. This can be seen through our hypothesis testing, which showed a significant difference in the average prevalence of obesity between the years 2015 and 2012, and then again between 2012 and previous years. The prevalence of obesity/overweight is not normally distributed, however, all regions exhibit a steady increase with years. Moreover, it is observed that women seem to have a higher prevalence of obesity in all regions across the globe, but not necessarily in overweight prevalence. Females have a higher mean BMI than males. The mean BMI of females is about 25, but the mean BMI of male is about 24.

In Africa there has been a noticeable steeper increase of overweight and obesity prevalence in females compared to that of males. In Europe it appears that the prevalence of overweight in males has increased more quickly than that of females. However, besides these two mentioned the other regions show similar rise of overweight prevalence over the time period for both sexes. Also, Eastern Mediterranean, and Europe have the highest mean BMI which are approximately 25. However, South-East Asia has the lowest mean BMI which is approximately 21. From our hypothesis testing it has been concluded that there is at least one region that has a statistically significant difference in the likelihood of death due to obesity related disease between the ages of 30 and 70. This shows that regions have a significant impact of obesity related deaths for individuals within this age group.

It is also important to understand that the obesity distribution peaks at a much lower prevalence than the overweight distribution. This is because all obese people are considered overweight but not all overweight people are considered obese. So naturally the

prevalence of overweight in our regions is going to be higher than prevalence of obesity because overweight is a larger “bucket” so to speak.

The next important conclusion is that European countries are more aware of their population’s habits and BMI health, while African countries are not as aware. This is supported by the fact that Africa had a noticeably large portion of countries that said they did not have surveys. However, European countries had a noticeably larger portion of the countries saying they did have surveys for all three categories in 2021. The exploratory analysis also helps in building a relation between surveying the obesity and overweight trends with factors that can cause obesity. A large number of countries that did have a survey on overweight/obesity in 2021, also had a survey on unhealthy diets and physical inactivity. This could suggest that countries that are surveying are trying to get a holistic perspective of their populations.