TITLE: GLOBAL DEMAND FOR ORGANIC FOOD

1 INTRODUCTION

1.1Problem description

People tend to adopt organic food habbit nowadays. Is this transition mainly because of health concerns or because of the premium prices or the export/import quality of the food.

1.2 Questions

- What type of customers (Demographic attributes,age group) are interested towards organic food products?
- How much a country invests for the organic food? (this question is answered in terms of land used for organic farming for UK)
- From the above question,the states where most of the exports for organic food products happen can be found.
- Among the organic foods, which is of greater demand(like fruits,vegetables,dairy products,meat)?
- How does the price of organic food vary with the conventional food?

1.3 Motivation

The demand for organic food have grown dramatically in the last decade despites its premium prices. This visualisation aims to study about the intention customers and how does a country contribute to its organic sales.[2]

2 Data Wrangling:

2.1 Data Source:

- UK Customers: To study about the customers buying organic food in the United Kingdom (Tabular data : 22,000 observations, 18 variables)

- USA Customers To study about the demographic characteristics of customers buying organic food in the United States(Statistical Data from a research paper in pdf format: 20 observations, 3 Variables)

- Organic-Products To study about the greater demand among the organic food(Data from a paper Report is pdf fomat: The data that has been recorded for four years 2004,2005,2006,2007,2008,2009, for fruits,vegetables,grains,poultry and meat products including the number of total share,purchases made by the customers; Data Taken from the USA government site)

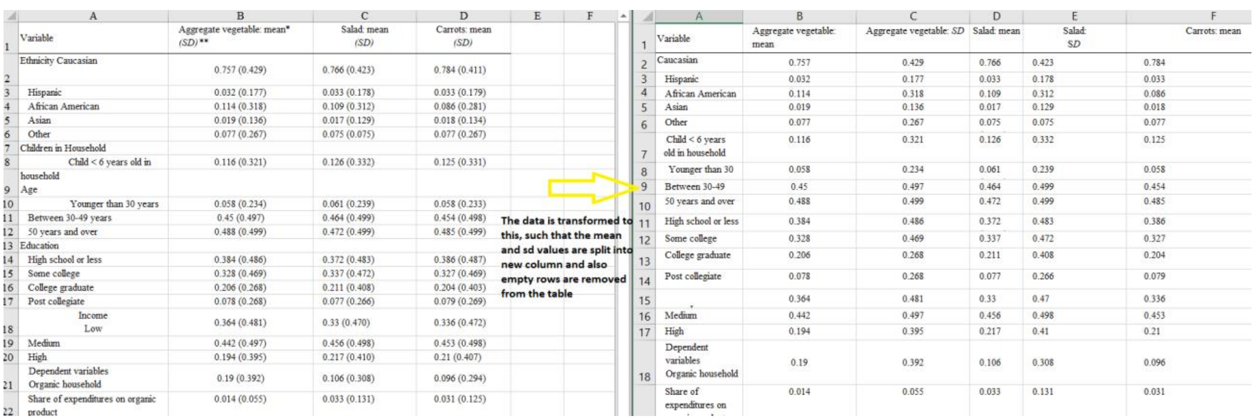
- Organic-prices To compare the prices of organic food products and conventional food(Tabular Data: The data that has been recorded for four years 2010,20122012,2013 for each month for fruits,vegetables,grains,poultry and meat products;Data Taken from the USA government site)

- USA-FARMING To study about how a country(USA) invests for its organic production (Tabular Data: Data from the USA government site is taken which records the for the years 2011-1997, for pastures,cropfield,livestock. Data in Excel format, 28 observations(each of the state) for the years 1997-2011)

- UK-FARMING To study about how a country(UK) invests for its organic production (Tabular Data: Data from the UK government site is taken which records the for the years 2002-2017)

2.2Data Transformation:

- For the USA Customers dataset, the research paper in pdf format is converted to excel using online too smallpdf, and then from the excel the table is extracted



	A	B	C	D	E	F
1	Variable	Aggregate vegetable: mean* (SD)**	Salad: mean (SD)	Carrots: mean (SD)		
2	Ethnicity Caucasian	0.757 (0.429)	0.766 (0.423)	0.784 (0.411)		
3	Hispanic	0.032 (0.177)	0.033 (0.178)	0.033 (0.179)		
4	African American	0.114 (0.318)	0.109 (0.312)	0.086 (0.281)		
5	Asian	0.019 (0.136)	0.017 (0.129)	0.018 (0.134)		
6	Other	0.077 (0.267)	0.075 (0.075)	0.077 (0.267)		
7	Children in Household					
8	Child < 6 years old in	0.116 (0.321)	0.126 (0.332)	0.125 (0.331)		
9	household Age					
10	Younger than 30 years	0.058 (0.234)	0.061 (0.239)	0.058 (0.233)		
11	Between 30-49 years	0.45 (0.497)	0.464 (0.499)	0.454 (0.498)		
12	50 years and over	0.488 (0.499)	0.472 (0.499)	0.485 (0.499)		
13	Education					
14	High school or less	0.384 (0.486)	0.372 (0.483)	0.386 (0.487)		
15	Some college	0.328 (0.469)	0.337 (0.472)	0.327 (0.469)		
16	College graduate	0.206 (0.268)	0.211 (0.408)	0.204 (0.403)		
17	Post collegiate	0.078 (0.268)	0.077 (0.266)	0.079 (0.269)		
18	Income Low	0.364 (0.481)	0.33 (0.470)	0.336 (0.472)		
19	Medium	0.442 (0.497)	0.456 (0.498)	0.453 (0.498)		
20	High	0.194 (0.395)	0.217 (0.410)	0.21 (0.407)		
21	Dependent variables Organic household	0.19 (0.392)	0.106 (0.308)	0.096 (0.294)		
22	Share of expenditures on organic product	0.014 (0.055)	0.033 (0.131)	0.031 (0.125)		

The data is transformed to this, such that the mean and sd values are split into new column and also empty rows are removed from the table

	A	B	C	D	E	F
1	Variable	Aggregate vegetable: mean	Aggregate vegetable: *SD*	Salad: mean	Salad: *SD*	Carrots: mean
2	Caucasian	0.757	0.429	0.766	0.423	0.784
3	Hispanic	0.032	0.177	0.033	0.178	0.033
4	African American	0.114	0.318	0.109	0.312	0.086
5	Asian	0.019	0.136	0.017	0.129	0.018
6	Other	0.077	0.267	0.075	0.075	0.077
7	Child < 6 years old in household	0.116	0.321	0.126	0.332	0.125
8	Younger than 30	0.058	0.234	0.061	0.239	0.058
9	Between 30-49	0.45	0.497	0.464	0.499	0.454
10	50 years and over	0.488	0.499	0.472	0.499	0.485
11	High school or less	0.384	0.486	0.372	0.483	0.386
12	Some college	0.328	0.469	0.337	0.472	0.327
13	College graduate	0.206	0.268	0.211	0.408	0.204
14	Post collegiate	0.078	0.268	0.077	0.266	0.079
15		0.364	0.481	0.33	0.47	0.336
16	Medium	0.442	0.497	0.456	0.498	0.453
17	High	0.194	0.395	0.217	0.41	0.21
18	Dependent variables Organic household	0.19	0.392	0.106	0.308	0.096
	Share of expenditures on	0.014	0.055	0.033	0.131	0.031

- For the Organic-Products dataset, the research paper in pdf format is converted to excel using online too smallpdf, and then from the excel the table is extracted.

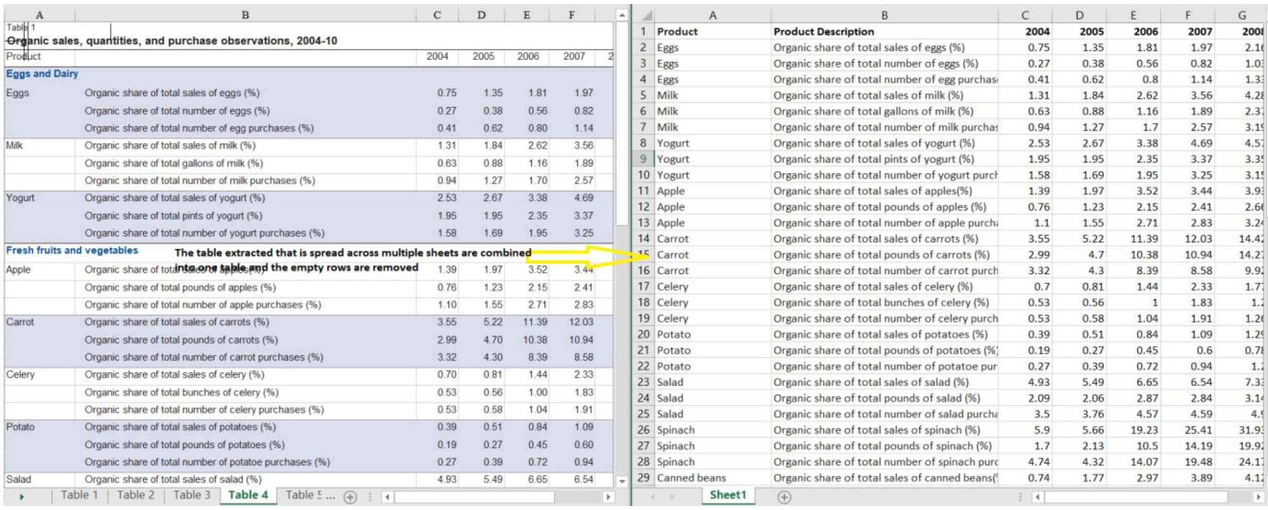


Table 1

Organic sales, quantities, and purchase observations, 2004-10

Product		2004	2005	2006	2007
Eggs and Dairy					
Eggs	Organic share of total sales of eggs (%)	0.75	1.35	1.81	1.97
	Organic share of total number of eggs (%)	0.27	0.38	0.56	0.82
	Organic share of total number of egg purchases (%)	0.41	0.62	0.80	1.14
Milk	Organic share of total sales of milk (%)	1.31	1.84	2.62	3.56
	Organic share of total gallons of milk (%)	0.63	0.88	1.16	1.89
	Organic share of total number of milk purchases (%)	0.94	1.27	1.70	2.57
Yogurt	Organic share of total sales of yogurt (%)	2.53	2.67	3.38	4.69
	Organic share of total pints of yogurt (%)	1.95	1.95	2.35	3.37
	Organic share of total number of yogurt purchases (%)	1.58	1.69	1.95	3.25
Fresh fruits and vegetables					
Apple	Organic share of total sales of apples (%)	1.39	1.97	3.52	3.44
	Organic share of total pounds of apples (%)	0.76	1.23	2.15	2.41
	Organic share of total number of apple purchases (%)	1.10	1.55	2.71	2.83
Carrot	Organic share of total sales of carrots (%)	3.55	5.22	11.39	12.03
	Organic share of total pounds of carrots (%)	2.99	4.70	10.38	10.94
	Organic share of total number of carrot purchases (%)	3.32	4.30	8.39	8.58
Celery	Organic share of total sales of celery (%)	0.70	0.81	1.44	2.33
	Organic share of total bunches of celery (%)	0.53	0.56	1.00	1.83
	Organic share of total number of celery purchases (%)	0.53	0.58	1.04	1.91
Potato	Organic share of total sales of potatoes (%)	0.39	0.51	0.84	1.09
	Organic share of total pounds of potatoes (%)	0.19	0.27	0.45	0.60
	Organic share of total number of potatoe purchases (%)	0.27	0.39	0.72	0.94
Salad	Organic share of total sales of salad (%)	4.93	5.49	6.65	6.54

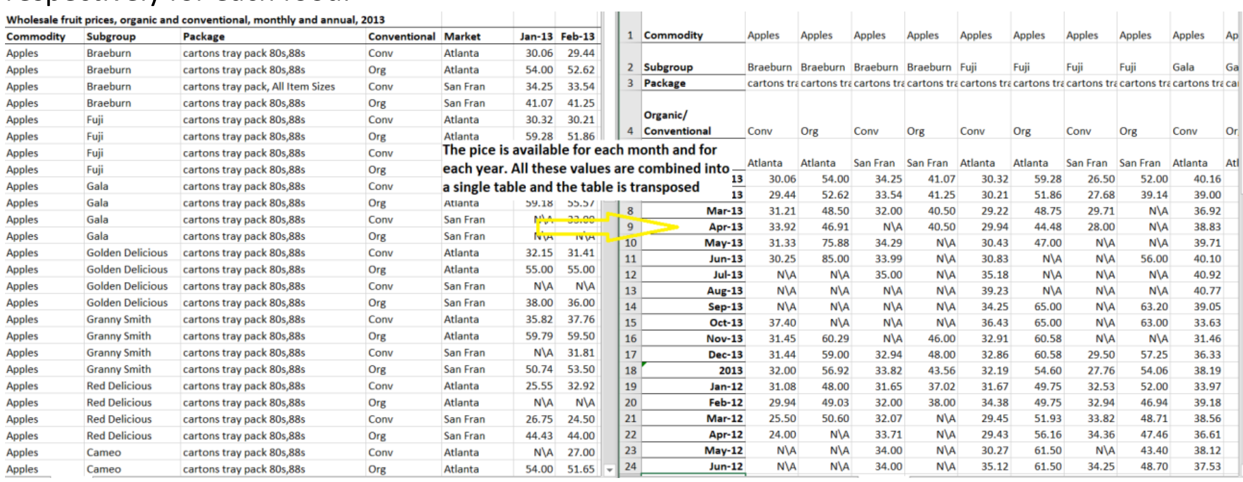
Table 1 | Table 2 | Table 3 | Table 4 | Table 5 ...

The table extracted that is spread across multiple sheets are combined into one table and the empty rows are removed

	Product	Product Description	2004	2005	2006	2007	2008
2	Eggs	Organic share of total sales of eggs (%)	0.75	1.35	1.81	1.97	2.1
3	Eggs	Organic share of total number of eggs (%)	0.27	0.38	0.56	0.82	1.0
4	Eggs	Organic share of total number of egg purchas	0.41	0.62	0.8	1.14	1.3
5	Milk	Organic share of total sales of milk (%)	1.31	1.84	2.62	3.56	4.2
6	Milk	Organic share of total gallons of milk (%)	0.63	0.88	1.16	1.89	2.3
7	Milk	Organic share of total number of milk purchas	0.94	1.27	1.7	2.57	3.1
8	Yogurt	Organic share of total sales of yogurt (%)	2.53	2.67	3.38	4.69	4.5
9	Yogurt	Organic share of total pints of yogurt (%)	1.95	1.95	2.35	3.37	3.3
10	Yogurt	Organic share of total number of yogurt purch	1.58	1.69	1.95	3.25	3.1
11	Apple	Organic share of total sales of apples(%)	1.39	1.97	3.52	3.44	3.9
12	Apple	Organic share of total pounds of apples (%)	0.76	1.23	2.15	2.41	2.6
13	Apple	Organic share of total number of apple purch	1.1	1.55	2.71	2.83	3.2
14	Carrot	Organic share of total sales of carrots (%)	3.55	5.22	11.39	12.03	14.4
15	Carrot	Organic share of total pounds of carrots (%)	2.99	4.7	10.38	10.94	14.2
16	Carrot	Organic share of total number of carrot purch	3.32	4.3	8.39	8.58	9.9
17	Celery	Organic share of total sales of celery (%)	0.7	0.81	1.44	2.33	1.7
18	Celery	Organic share of total bunches of celery (%)	0.53	0.56	1	1.83	1.
19	Celery	Organic share of total number of celery purch	0.53	0.58	1.04	1.91	1.2
20	Potato	Organic share of total sales of potatoes (%)	0.39	0.51	0.84	1.09	1.2
21	Potato	Organic share of total pounds of potatoes (%)	0.19	0.27	0.45	0.6	0.7
22	Potato	Organic share of total number of potatoe pur	0.27	0.39	0.72	0.94	1.
23	Salad	Organic share of total sales of salad (%)	4.93	5.49	6.65	6.54	7.3
24	Salad	Organic share of total pounds of salad (%)	2.09	2.06	2.87	2.84	3.1
25	Salad	Organic share of total number of salad purcha	3.5	3.76	4.57	4.59	4.
26	Spinach	Organic share of total sales of spinach (%)	5.9	5.66	19.23	25.41	31.9
27	Spinach	Organic share of total pounds of spinach (%)	1.7	2.13	10.5	14.19	19.9
28	Spinach	Organic share of total number of spinach purc	4.74	4.32	14.07	19.48	24.1
29	Canned beans	Organic share of total sales of canned beans(	0.74	1.77	2.97	3.89	4.1

Sheet1

- For the Organic-prices dataset, the data is available for fruits,vegetables,grains,poultry and meat for each month and for each year. All the year values are combined into a single table respectively for each food.



Wholesale fruit prices, organic and conventional, monthly and annual, 2013

Commodity	Subgroup	Package	Conventional	Market	Jan-13	Feb-13
Apples	Braeburn	cartons tray pack 80s,88s	Conv	Atlanta	30.06	29.44
Apples	Braeburn	cartons tray pack 80s,88s	Org	Atlanta	54.00	52.62
Apples	Braeburn	cartons tray pack, All Item Sizes	Conv	San Fran	34.25	33.54
Apples	Braeburn	cartons tray pack 80s,88s	Org	San Fran	41.07	41.25
Apples	Fuji	cartons tray pack 80s,88s	Conv	Atlanta	30.32	30.21
Apples	Fuji	cartons tray pack 80s,88s	Org	Atlanta	59.28	51.86
Apples	Fuji	cartons tray pack 80s,88s	Conv			
Apples	Fuji	cartons tray pack 80s,88s	Org			
Apples	Gala	cartons tray pack 80s,88s	Conv			
Apples	Gala	cartons tray pack 80s,88s	Org	Atlanta	59.18	55.57
Apples	Gala	cartons tray pack 80s,88s	Conv	San Fran	N\A	33.00
Apples	Gala	cartons tray pack 80s,88s	Org	San Fran	N\A	N\A
Apples	Golden Delicious	cartons tray pack 80s,88s	Conv	Atlanta	32.15	31.41
Apples	Golden Delicious	cartons tray pack 80s,88s	Org	Atlanta	55.00	55.00
Apples	Golden Delicious	cartons tray pack 80s,88s	Conv	San Fran	N\A	N\A
Apples	Golden Delicious	cartons tray pack 80s,88s	Org	San Fran	38.00	36.00
Apples	Granny Smith	cartons tray pack 80s,88s	Conv	Atlanta	35.82	37.76
Apples	Granny Smith	cartons tray pack 80s,88s	Org	Atlanta	59.79	59.50
Apples	Granny Smith	cartons tray pack 80s,88s	Conv	San Fran	N\A	31.81
Apples	Granny Smith	cartons tray pack 80s,88s	Org	San Fran	50.74	53.50
Apples	Red Delicious	cartons tray pack 80s,88s	Conv	Atlanta	25.55	32.92
Apples	Red Delicious	cartons tray pack 80s,88s	Org	Atlanta	N\A	N\A
Apples	Red Delicious	cartons tray pack 80s,88s	Conv	San Fran	26.75	24.50
Apples	Red Delicious	cartons tray pack 80s,88s	Org	San Fran	44.43	44.00
Apples	Cameo	cartons tray pack 80s,88s	Conv	Atlanta	N\A	27.00
Apples	Cameo	cartons tray pack 80s,88s	Org	Atlanta	54.00	51.65

The pice is available for each month and for each year. All these values are combined into a single table and the table is transposed

1	Commodity	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Ap
2	Subgroup	Braeburn	Braeburn	Braeburn	Braeburn	Fuji	Fuji	Fuji	Fuji	Gala	Ga
3	Package	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	ca
4	Organic/ Conventional	Conv	Org	Conv	Org	Conv	Org	Conv	Org	Conv	Or
		Atlanta	Atlanta	San Fran	San Fran	Atlanta	Atlanta	San Fran	San Fran	Atlanta	Atl
	13	30.06	54.00	34.25	41.07	30.32	59.28	26.50	52.00	40.16	
	13	29.44	52.62	33.54	41.25	30.21	51.86	27.68	39.14	39.00	
8	Mar-13	31.21	48.50	32.00	40.50	29.22	48.75	29.71	N\A	36.92	
9	Apr-13	33.92	46.91	N\A	40.50	29.94	44.48	28.00	N\A	38.83	
10	May-13	31.33	75.88	34.29	N\A	30.43	47.00	N\A	N\A	39.71	
11	Jun-13	30.25	85.00	33.99	N\A	30.83	N\A	N\A	56.00	40.10	
12	Jul-13	N\A	N\A	35.00	N\A	35.18	N\A	N\A	N\A	40.92	
13	Aug-13	N\A	N\A	N\A	N\A	39.23	N\A	N\A	N\A	40.77	
14	Sep-13	N\A	N\A	N\A	N\A	34.25	65.00	N\A	63.20	39.05	
15	Oct-13	37.40	N\A	N\A	N\A	36.43	65.00	N\A	63.00	33.63	
16	Nov-13	31.45	60.29	N\A	46.00	32.91	60.58	N\A	N\A	31.46	
17	Dec-13	31.44	59.00	32.94	48.00	32.86	60.58	29.50	57.25	36.33	
18	2013	32.00	56.92	33.82	43.56	32.19	54.60	27.76	54.06	38.19	
19	Jan-12	31.08	48.00	31.65	37.02	31.67	49.75	32.53	52.00	33.97	
20	Feb-12	29.94	49.03	32.00	38.00	34.38	49.75	32.94	46.94	39.18	
21	Mar-12	25.50	50.60	32.07	N\A	29.45	51.93	33.82	48.71	38.56	
22	Apr-12	24.00	N\A	33.71	N\A	29.43	56.16	34.36	47.46	36.61	
23	May-12	N\A	N\A	34.00	N\A	30.27	61.50	N\A	43.40	38.12	
24	Jun-12	N\A	N\A	34.00	N\A	35.12	61.50	34.25	48.70	37.53	

3 DATA CLEANING:

For the UK Customers data set, using python the data cleaning task is done

Coverage Anomalies: Missing values are found and are then imputed.

In [2]:
```
import pandas as pd
#dates are read as strings so parsing those columns are date time object
df = pd.read_csv('qn1_Uk.csv',index_col=0,parse_dates= ['DOB','EDATE','LCDATE'])
df.head()
```

Parsing the Date columns as date field and not as string for manipulation

Out[2]:

	CUSTID	GENDER	DOB	EDATE	AGE	AGEGRP1	AGEGRP2	TV_REG	NGROUP	NEIGHBORHOOD	LCDATE	ORGANICS	BILL	REGION	CLASS	OF
1	140	U	1921-09-16	1998-02-23	76.0	60-80	70-80	Wales & West	C	16.0	1994-11-07	0	16000.00	Midlands	Gold	
2	620	U	1949-02-12	1998-02-23	49.0	40-60	40-50	Wales & West	D	35.0	1993-06-04	0	6000.00	Midlands	Gold	
3	868	F	1927-11-27	1998-02-23	70.0	60-80	70-80	Wales & West	D	27.0	1990-08-02	1	0.02	Midlands	Silver	
4	1120	M	1932-04-10	1998-02-23	65.0	60-80	60-70	Midlands	F	51.0	1991-07-01	1	0.01	Midlands	Tin	
5	2313	F	1929-05-21	1998-02-23	68.0	60-80	60-70	Midlands	A	4.0	1990-03-01	0	0.01	Midlands	Tin	

Here AGEGRP1 and AGEGRP2 does imply the same meaning, so the AGEGRP2 Column is dropped from the table

Number of rows and columns

In [3]:
```
print("Rows: ",len(df))
print("\nColumns: ",len(df.columns))
```

```
Rows:  22223

Columns:  18
```

Identifying missing values

Finding the those age groups buying organic food is the subject of interest. Here Age column has null values

In [4]:
```
df.columns[df.isnull().any()]
```

Out[4]:
```
Index(['GENDER', 'AGE', 'AGEGRP1', 'AGEGRP2', 'TV_REG', 'NGROUP',
       'NEIGHBORHOOD', 'LCDATE', 'REGION', 'AFFL', 'LTIME'],
      dtype='object')
```

seperating the year date and month from the dob field, so that the missing age values can be imputed from the dob' year value to find the age

In [16]:
```
df['dob_year'] = df['DOB'].dt.year
df.head(2)
```
Extracting the year from the DOB Field

Out[16]:

	CUSTID	GENDER	DOB	EDATE	AGE	AGEGRP1	TV_REG	NGROUP	NEIGHBORHOOD	LCDATE	ORGANICS	BILL	REGION	CLASS	ORGYN	AFFL
1	140	U	1921-09-16	1998-02-23	76.0	60-80	Wales & West	C	16.0	1994-11-07	0	16000.0	Midlands	Gold	0	10.0
2	620	U	1949-02-12	1998-02-23	49.0	40-60	Wales & West	D	35.0	1993-06-04	0	6000.0	Midlands	Gold	0	4.0

In [17]:
```
# creating bool series True for NaN values for the age columns since we are intrersted in the age group of the customers
bool_series = pd.isnull(df["AGE"])
print(len(df[bool_series]))
df[bool_series].head(2)
```
```
1508
```
Those Nan value from the AGE column are stored as a boolean variables

Out[17]:

	CUSTID	GENDER	DOB	EDATE	AGE	AGEGRP1	TV_REG	NGROUP	NEIGHBORHOOD	LCDATE	ORGANICS	BILL	REGION	CLASS	ORGYN	AFFL
12	9814	M	1942-12-31	1998-02-23	NaN	NaN	London	C	24.0	1997-11-10	1	5000.00	South East	Silver	1	5.0
19	15350	F	1966-11-30	1998-02-23	NaN	NaN	C Scotland	E	40.0	1993-09-22	1	0.01	Scottish	Tin	1	7.0

The Age column is imputed with values by finding the age using the extracted field

In [21]:
```
df[bool_series] =df[bool_series].apply(lambda x: x.fillna(1998-df['dob_year']),axis=0)#finding the age for all the missings rec
df[bool_series]
```

Out[21]:

	CUSTID	GENDER	DOB	EDATE	AGE	AGEGRP1	TV_REG	NGROUP	NEIGHBORHOOD	LCDATE	ORGANICS	BILL	REGION	CLASS	OR
12	9814	M	1942-12-31	1998-02-23	56.0	56	London	C	24.0	1997-11-10 00:00:00	1	5000.00	South East	Silver	
19	15350	F	1966-11-30	1998-02-23	32.0	32	C Scotland	E	40.0	1993-09-22 00:00:00	1	0.01	Scottish	Tin	
34	39330	M	1967-12-06	1998-02-23	31.0	31	East	F	50.0	1994-04-22 00:00:00	0	0.01	Midlands	Tin	
58	83852	F	1937-02-28	1998-02-23	61.0	61	London	A	4.0	1994-10-27 00:00:00	0	5000.00	South East	Silver	
78	126586	F	1920-08-03	1998-02-23	78.0	78	Midlands	C	18.0	1997-11-06 00:00:00	0	15000.00	Midlands	Gold	

In [27]:
```
#for every value in age, if thh condition is checked the value in agegroup1 is replace
df['AGEGRP1'] = df['AGE'].apply(lambda s: '<20'  if(s < 20.0) else ('60-80' if s<80.0 and s>=60.0 else ('20-40' if s<40.0 and s>
```
The AGEGRP1 Values are imputed based on a condition check from the AGE column

In [26]:
```
df.AGEGRP1.unique()
```
Out[26]:
```
array(['60-80', '40-60', '20-40', '<20', '80-100'], dtype=object)
```
There are no missing value in AGEGRP1 , thus for visualisation the columns that are required are imputed

Now that all of the Null values of AGE are imputed

In [17]:
```
df.AGEGRP1.value_counts()
```
Out[17]:
```
40-60     10431
60-80      8078
20-40      3692
<20          16
80-100        6
Name: AGEGRP1, dtype: int64
```

In [28]:
```
export_csv = df.to_excel ('uk_Supermarket.xlsx') #exporting the dataset to excel format
```
Exporting the dataframe to excel sheet

Coverage Anomalies: For the Organic-prices dataset, all the null values as changed to N/A values.

Commodity	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples	Apples
Subgroup	Braeburn	Braeburn	Braeburn	Braeburn	Fuji	Fuji	Fuji	Fuji	Gala	Gala	Gala	Gala	Golden De	Golden De	Golden De	Golden De	Granny Sm	Granny Sm	Granny Sm	Granny Sm	Red Delicic
Package	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra	cartons tra
Organic/ Conventional	Conv	Org	Conv	Org	Conv	Org	Conv	Org	Conv	Org	Conv	Org	Conv	Org	Conv	Org	Conv	Org	Conv	Org	Conv
Terminal Market	Atlanta	Atlanta	San Fran	San Fran	Atlanta	Atlanta	San Fran	San Fran	Atlanta	Atlanta	San Fran	San Fran	Atlanta	Atlanta	San Fran	San Fran	Atlanta	Atlanta	San Fran	San Fran	Atlanta
Jan-13	30.06	54.00	34.25	41.07	30.32	59.28	26.50	52.00	40.16	59.18	N\A	N\A	32.15	55.00	N\A	38.00	35.82	59.79	N\A	50.74	25.55
Feb-13	29.44	52.62	33.54	41.25	30.21	51.86	27.68	39.14	39.00	55.57	33.00	N\A	31.41	55.00	N\A	36.00	37.76	59.50	31.81	53.50	32.92
Mar-13	31.21	48.50	32.00	40.50	29.22	48.75	29.71	N\A	36.92	59.50	N\A	53.30	31.15	53.13	N\A	48.50	34.60	59.50	31.00	N\A	31.19
Apr-13	33.92	46.91	N\A	40.50	29.94	44.48	28.00	N\A	38.83	56.20	N\A	53.00	32.69	48.59	N\A	N\A	34.84	59.50	N\A	N\A	30.07
May-13	31.33	75.88	34.29	N\A	30.43	47.00	N\A	N\A	39.71	56.50	34.00	55.00	32.35	46.04	N\A	N\A	36.28	N\A	43.50	N\A	29.00

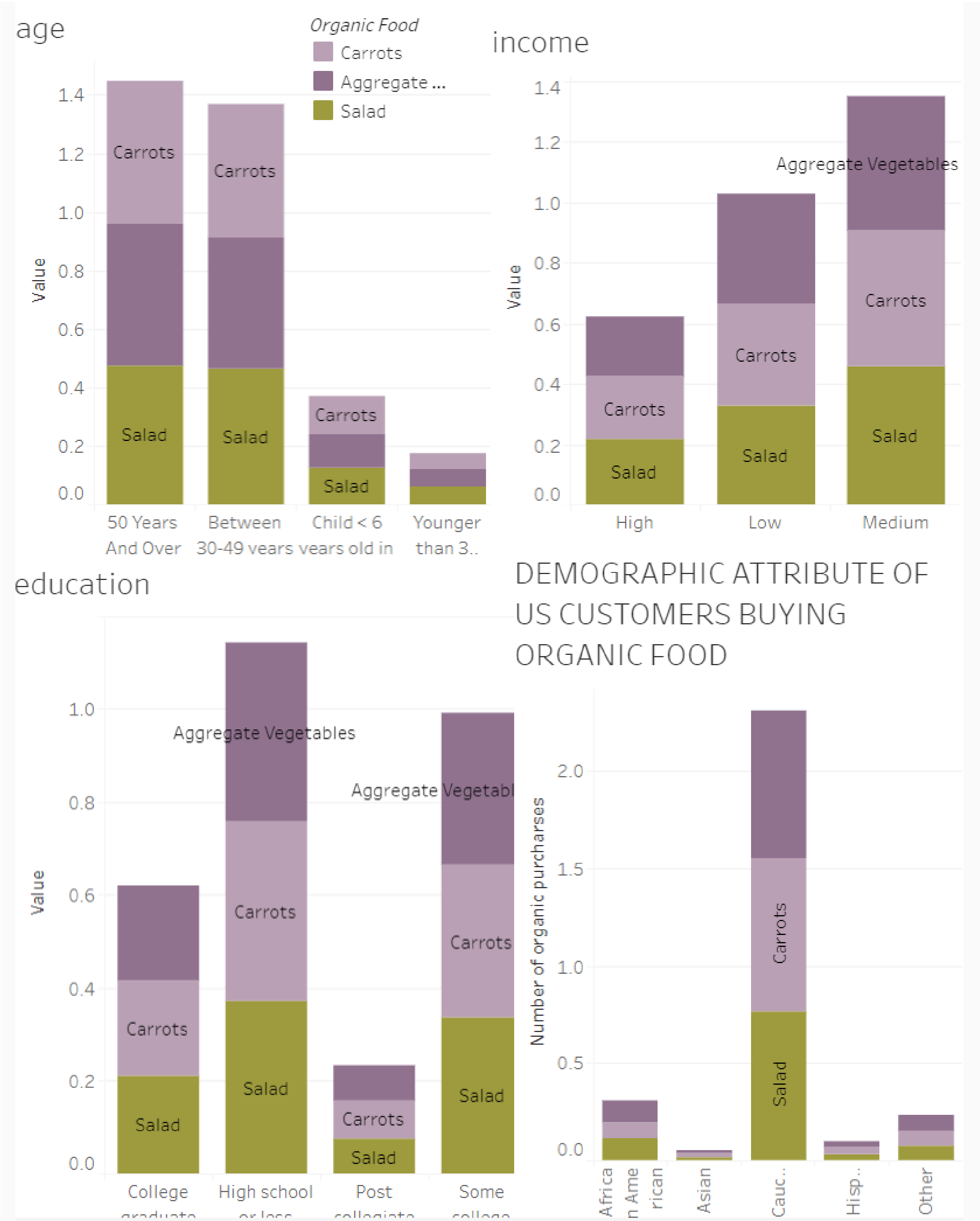
Each of the remaining data set . boundary check to identify Syntactic anomalies is made using Excel’s filter option.

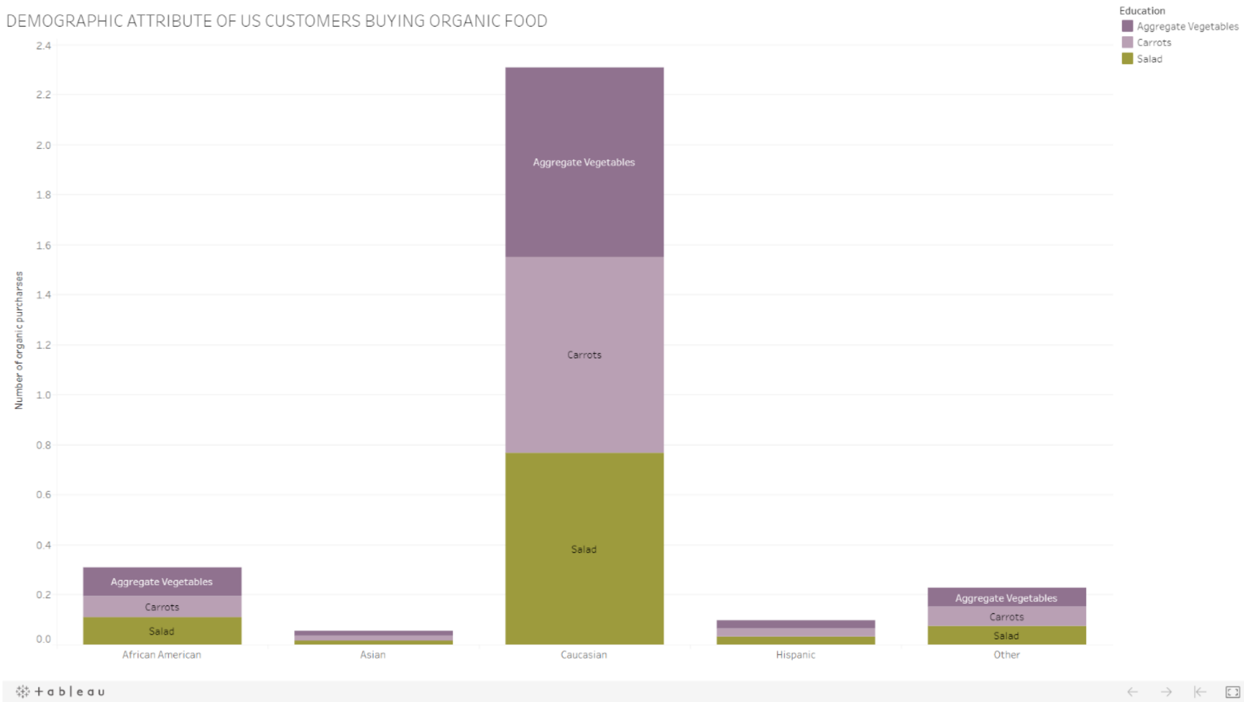
Also the null values are also fould using the filter option.

4. DATA EXPLORATION

TOOL USED: TABLEAU

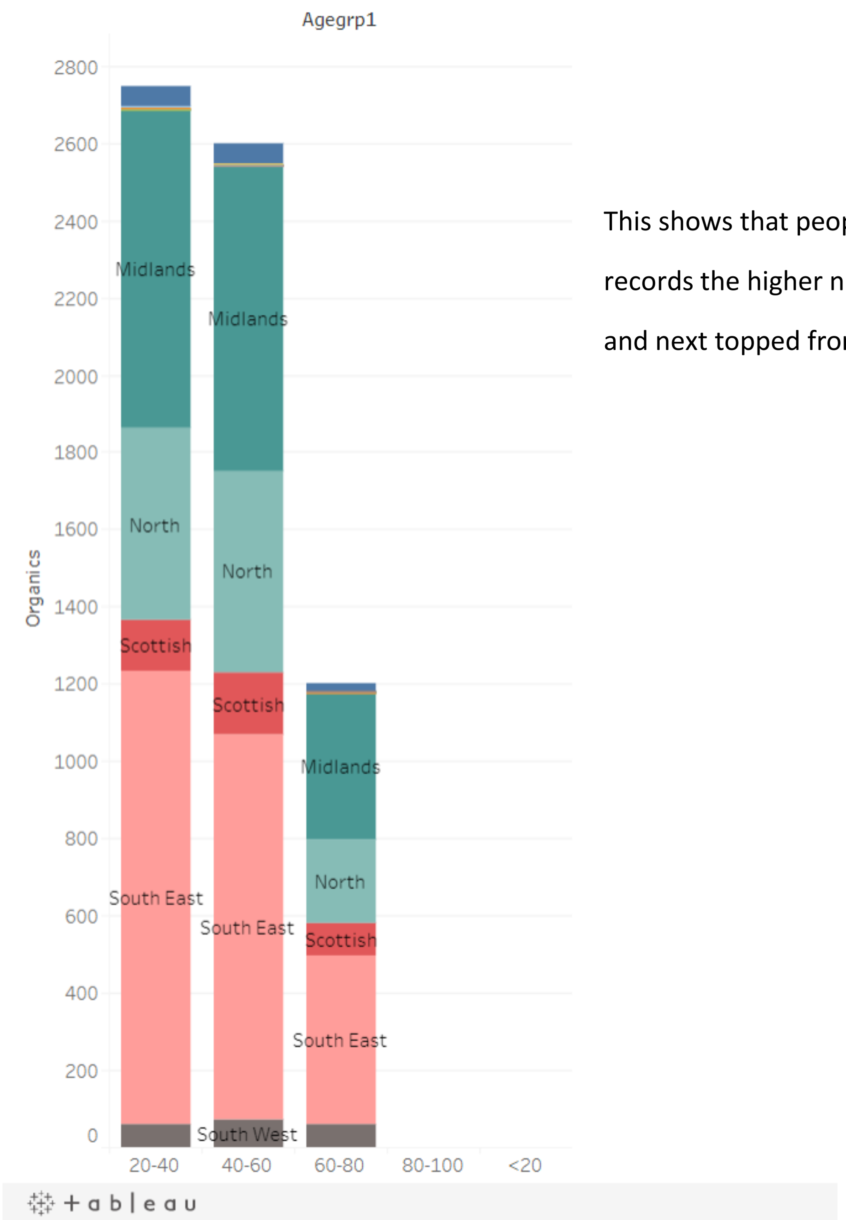
To study the demographic characteristics of people buying organic food, the statistics from the Nielsen Homescan data for the year 2004 is used. While observing,the Caucasian race records the higher number of purchases compared to the rest.


age
Organic Food
Carrots
Aggregate ...
Salad
Value
50 Years And Over
Between 30-49 years
Child < 6 years old in
Younger than 3..
income
Aggregate Vegetables
High
Low
Medium
education
Aggregate Vegetables
College
High school
Post
Some
DEMOGRAPHIC ATTRIBUTE OF US CUSTOMERS BUYING ORGANIC FOOD
Number of organic purcharses
Africa n Ame rican
Asian
Cauc..
Hisp..
Other



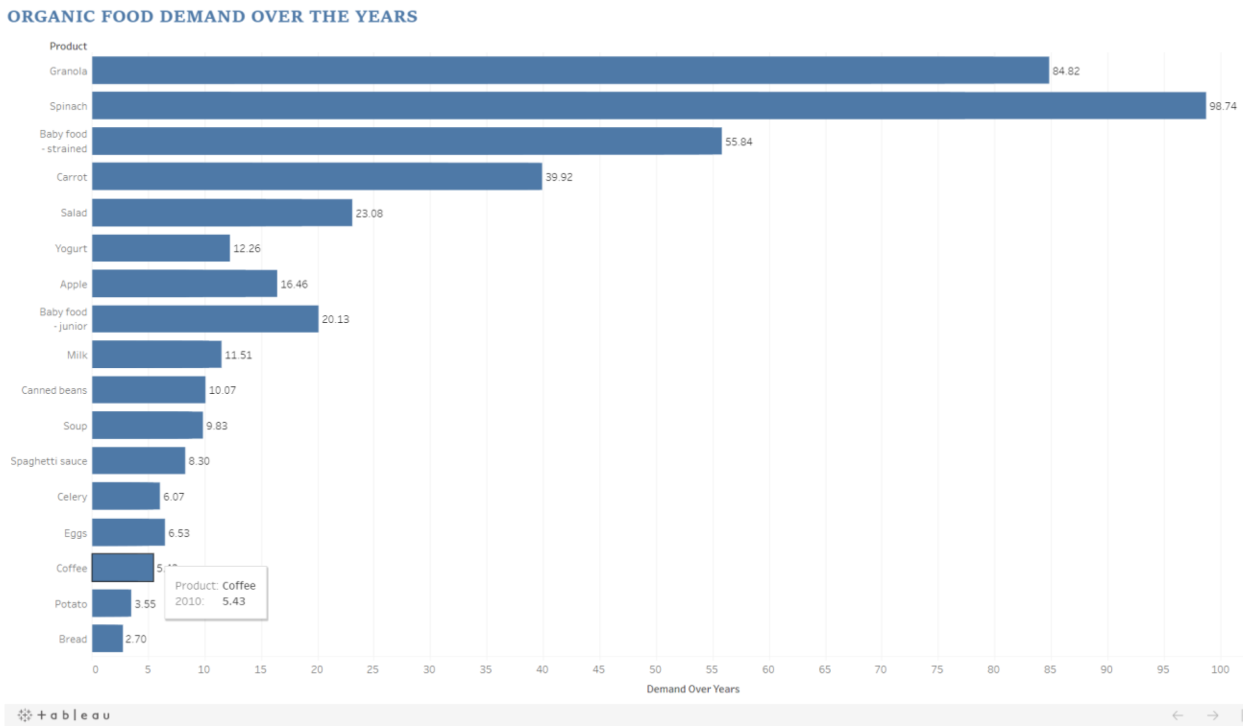
To study more about the Caucasian race,the following chart represents people who buy organic food in a supermarket from different region across the United Kingdom. Also to learn about which age group of people are interested in buying organic food.

UK CUSTOMERS INTERESTED IN ORGANIC FOOD

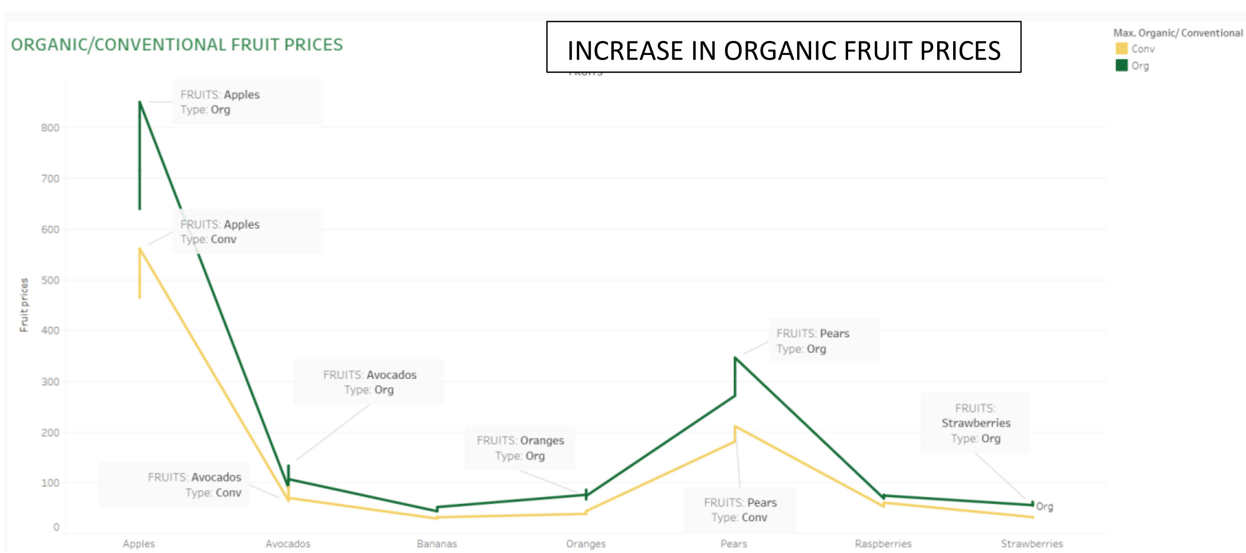


This shows that people between the age group 20 to 40 records the higher number of purchases for organic food and next topped from people of age group between 40 to 60

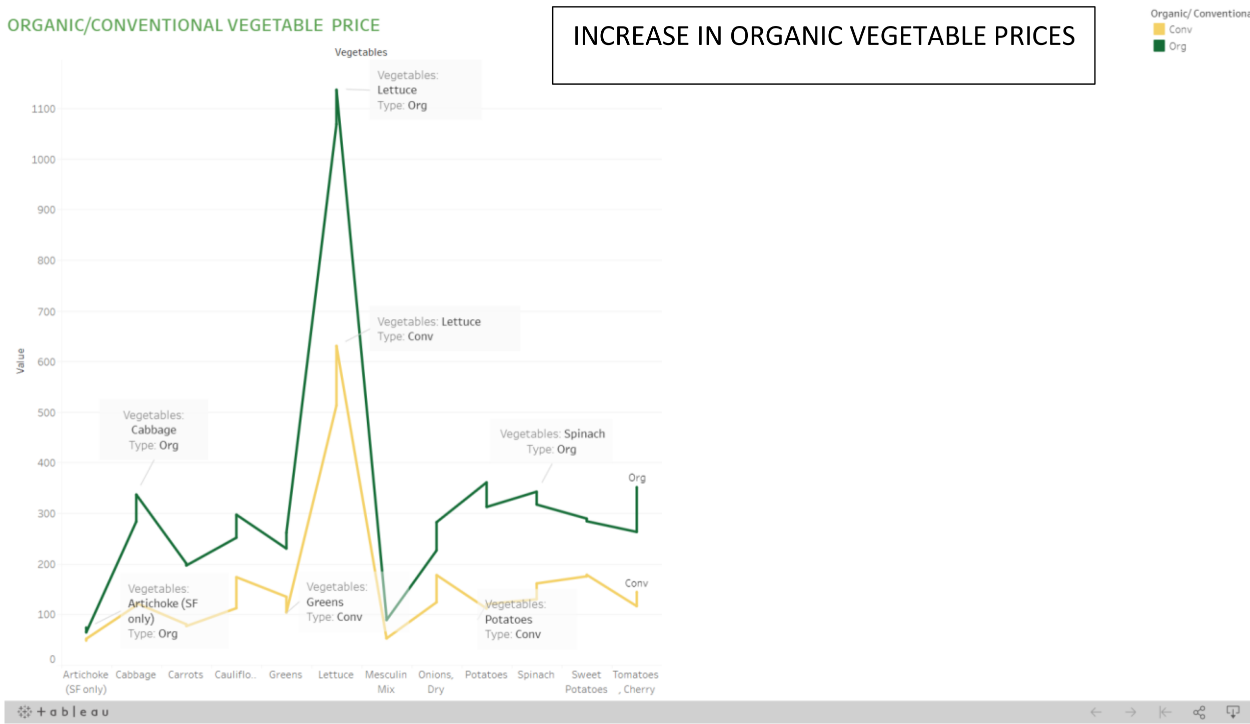
Among these organic food purchases the below chart potrays the demand for a particular organic food from the dataset recorded based on the total number of purchases from 2004 to 2010. The demand of organic vegetables ranks first intuitively.



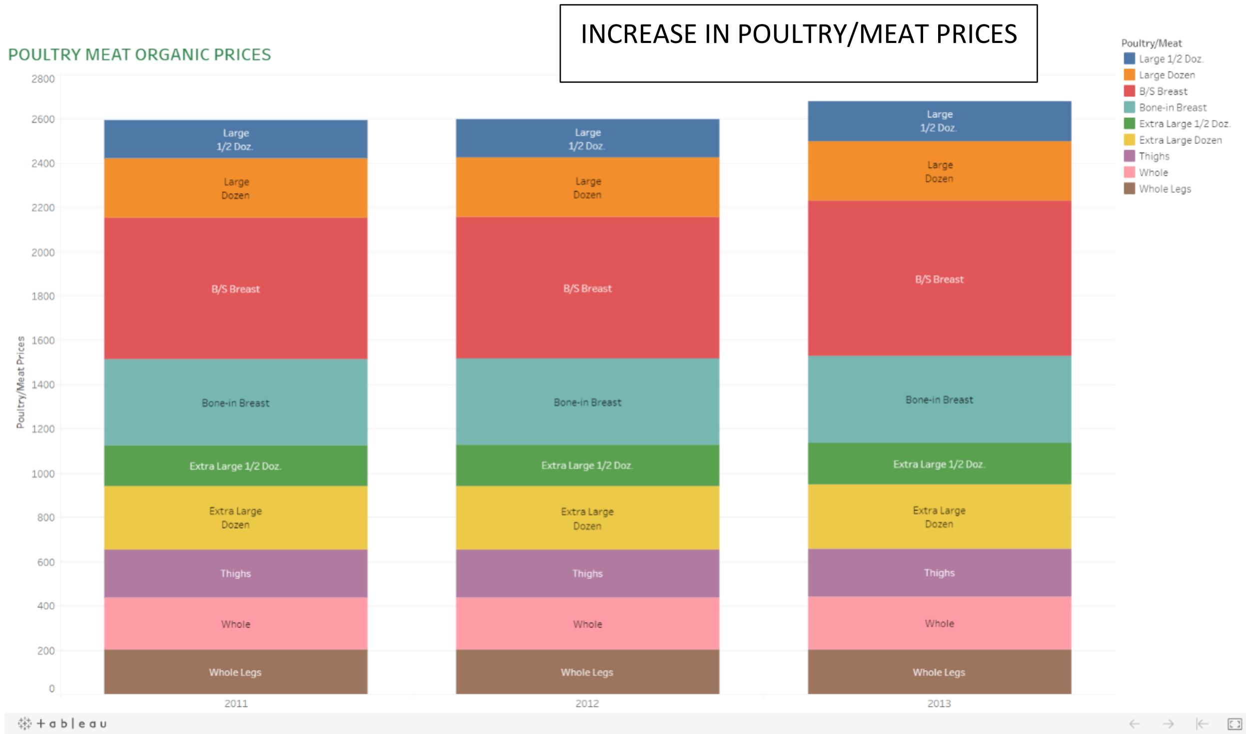
As the demand for organic food increases, the prices for organic food also increased compared to conventional food from the data that has been recorded for the years 2010 – 2013.The increase is observed for all type of food (i.e.)vegetable,fruits,grains,poultry and meat[1]



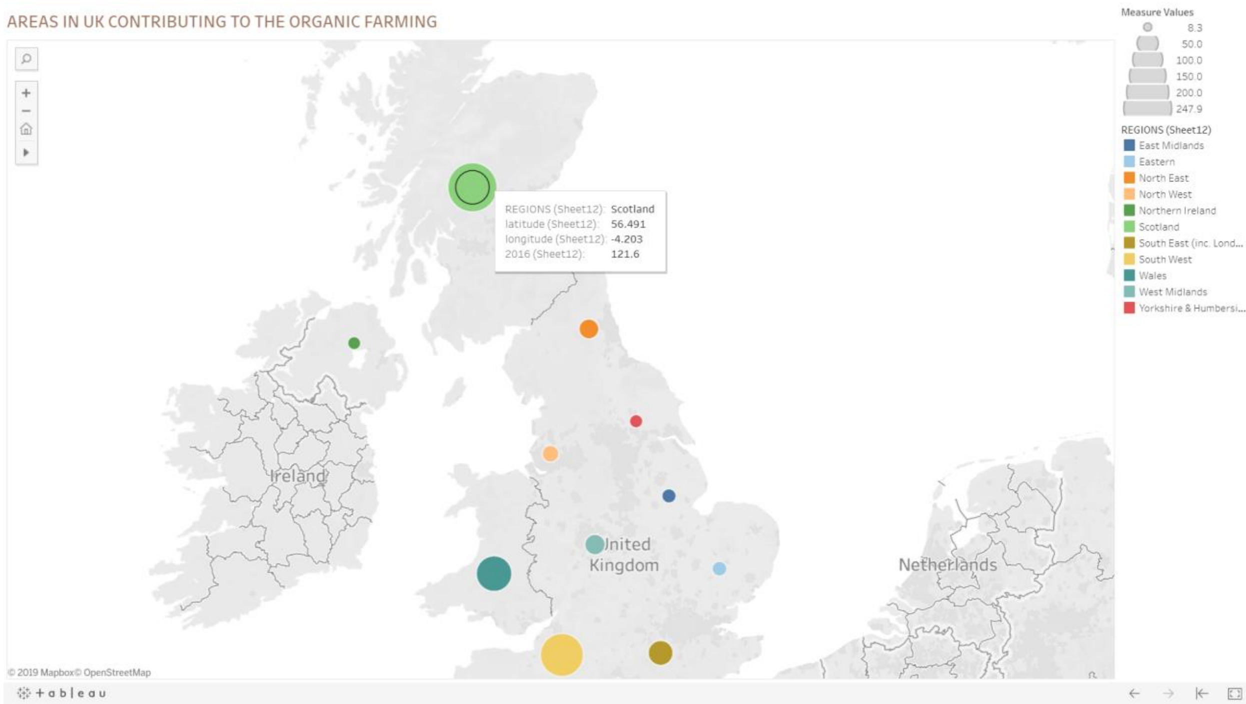
INCREASE IN ORGANIC VEGETABLE PRICES



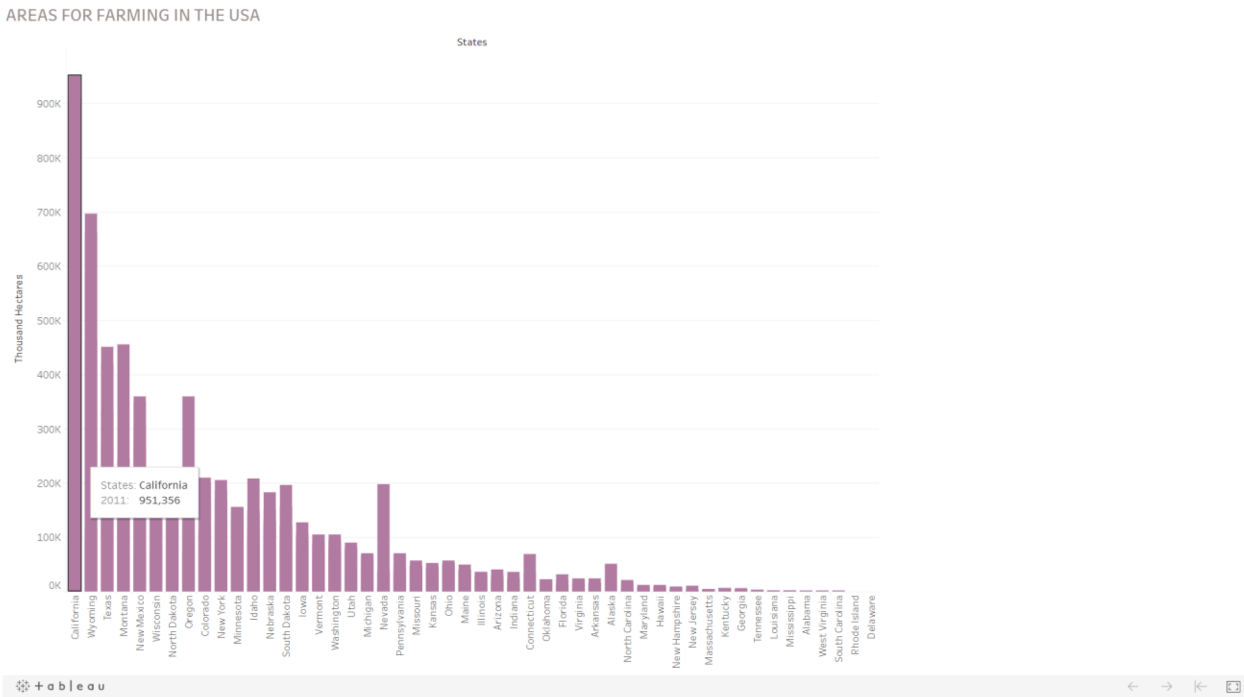
INCREASE IN POULTRY/MEAT PRICES



Both USA and UK customers tend to use a larger number of organic food. To visulaize how much a country investes for its organic food production the below map can be used.



In the United Kingdom,Scotland tend to cover a larger area for organic food production(in terms of crop fields and pastures). Then ranks the Wales and the South West UK.



In the USA, California ranks top among the areas allocated for organic farming. Thus from this it can be said that California and Scotland can be the major exports for the organic food products.

4 CONCLUSION

Thus from the above visualisations, we've learnt that

- The Caucasian race people tend to show a higher interest towards organic food products.
- Among those people who consume organic food, people from age group between 20-40 tend to have purchased larger number of organic foods.
- Organic vegetables are of greater demand when comapred to the rest of the food.
- As the demand for organic food products are growing every year, the number of purchases are increasing which inturns cause the number of prices for organic foods to increase as well.
- The areas for organic food production in USA and UK in terms of the area of field allocated for the organic food production is also visualized and also which among the states ranks highest in terms of area of field for either of the countries is also studied.
- The areas where the export of organic food products is higher is also visualized.

5 REFLECTION

- From this project; The Global demand for organic food products is identified by the number of purchases mades by customers in the two countrie taken under study (USA and UK).
- The demographic charactericteristics of these customers.
- Among he organic food, which is food is of demand is found from this project. The price variation for conventional food and organic food is studied due to the demand of organic food.
- At the end,which of the states in the USA and the UK has a greater area for organic food production and also the exporting area for organic food products is also observed from the visualisations produced for this project.

6 REFERENCES

[1] http://www.fao.org/organicag/oa-faq/oa-faq5/en/

[2] https://www.organicconsumers.org/news/demand-organic-food-growing-faster-domestic-supply