PROJECT 1: VISUALIZING DATA

PROJECT DATA: COFFEE PRODUCTION

```
coffee_production.head()
```

total_production	Angola	Bolivia (Plurinational State of)	Brazil	...	Venezuela	Viet Nam	Yemen
1990	50.345	122.777	27285.6286		1122.477	1310.288	0.0
1991	79.331	103.536	27293.4934		940.704	1437.848	0.0
1992	77.52	120.235	34603.3542	...	1215.298	2340.447	0.0
1993	32.608	50.823	28166.9786		1332.881	3020.216	0.0
1994	76.802	116.944	28192.047		988.996	3631.609	0.0

```
coffee_production.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 29 entries, 1990 to 2018
Data columns (total 56 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   Angola                            29 non-null     object
 1   Bolivia (Plurinational State of)  29 non-null     object
 2   Brazil                            29 non-null     object
 3   Burundi                           29 non-null     object
 4   Ecuador                           29 non-null     object
 5   Indonesia                         29 non-null     object
 6   Madagascar                        29 non-null     object

 ...

 52  Uganda                            29 non-null     object
 53  Venezuela                         29 non-null     object
 54  Viet Nam                          29 non-null     object
 55  Yemen                             29 non-null     object
dtypes: object(56)
memory usage: 12.9+ KB
```

MCG

PROJECT DATA: COFFEE IMPORTS

```
imports.head()
```

	imports	1990	1991	1992	1993	...	2011	2012	2013
0	Austria	1880.0	2058.0	2206.0	1836.0		1452.0	1559.0	1555.0
1	Belgium	NaN	NaN	NaN	NaN		5828.0	5668.0	5502.0
2	Belgium/Luxembourg	2015.0	1746.0	1828.0	2063.0	...	NaN	NaN	NaN
3	Bulgaria	268.0	200.0	182.0	397.0		482.0	560.0	609.0
4	Croatia	NaN	NaN	168.0	163.0		391.0	384.0	413.0

```
imports.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 147 entries, 0 to 146
Data columns (total 30 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   imports   147 non-null    object
 1   1990      138 non-null    float64
 2   1991      138 non-null    float64
 3   1992      144 non-null    float64
 4   1993      145 non-null    float64
 5   1994      145 non-null    float64

 ...

 27  2016      118 non-null    float64
 28  2017      118 non-null    float64
 29  2018      118 non-null    float64
dtypes: float64(29), object(1)
memory usage: 34.6+ KB
```

MCG

PROJECT DATA: COFFEE PRICES

```
prices.head()
```

	retail_prices	1990	1991	1992	...	2016	2017	2018
0	Austria	10.816777	10.088300	11.015453		12.450331	13.730684	14.635762
1	Cyprus	6.247241	6.181015	6.335541		11.699779	12.141280	12.781457
2	Denmark	8.410596	8.101545	8.366446	...	10.905077	11.103753	11.699779
3	Finland	6.578366	6.004415	5.430464		8.101545	9.050773	9.359823
4	France	8.233996	7.571744	5.099338		7.196468	7.505519	8.123620

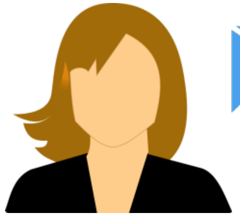
```
prices.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14 entries, 0 to 13
Data columns (total 30 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   retail_prices  14 non-null     object
 1   1990           14 non-null     float64
 2   1991           14 non-null     float64
 3   1992           14 non-null     float64
 4   1993           14 non-null     float64
 5   1994           14 non-null     float64

 ...

 27  2016           14 non-null     float64
 28  2017           14 non-null     float64
 29  2018           14 non-null     float64
dtypes: float64(29), object(1)
memory usage: 3.4+ KB
```

ASSIGNMENT: MID-COURSE PROJECT



NEW MESSAGE
September 2024 ,7

From: **Sarah Shark** *(Managing Director)*

Subject: **Coffee Industry Deep Dive**

Hi there,

I'm starting to trust you… which is rare. We just got an inquiry from a major coffee trader looking to get an outside view on the coffee industry. They're particularly interested in Brazil's production relative to other nations.

We'll also look at a comparison of importer volume vs the prices they pay to understand if we can unlock margin by diversifying into new markets.

Do well on this and you'll be on promotion track.

section03_coffee_project_part1.ipynb

Reply Forward

Key Objectives

1. Read in data from multiple csv files
2. Reshape the data to prepare it for visualization
3. Build & customize charts to communicate the key insights to the client

MCG

PROJECT 2: CUSTOM LAYOUTS

PROJECT DATA: OVERVIEW

Coffee Production

total_production	Angola	Bolivia (Plurinational State of)	Brazil	Burundi	Ecuador	Indonesia	Madagascar	...	Uganda	Venezuela	Viet Nam	Yemen
1990	50.345	122.777	27285.6286	487.393	1503.815	7441.383	982.447		1955.009	1122.477	1310.288	0.0
1991	79.331	103.536	27293.4934	667.199	2123.824	8493.196	932.513		2088.001	940.704	1437.848	0.0
1992	77.52	120.235	34603.3542	620.238	1185.48	5569.478	1121.684		2185.074	1215.298	2340.447	0.0
1993	32.608	50.823	28166.9786	393.354	2069.007	6743.288	441.859		3141.706	1332.881	3020.216	0.0
1994	76.802	116.944	28192.047	664.143	2375.766	5367.878	641.372		2392.753	988.996	3631.609	0.0

```
coffee_production.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 29 entries, 1990 to 2018
Data columns (total 56 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   Angola                            29 non-null     object
 1   Bolivia (Plurinational State of)  29 non-null     object
 2   Brazil                            29 non-null     object
 3   Burundi                           29 non-null     object
 4   Ecuador                           29 non-null     object
 5   Indonesia                         29 non-null     object
 6   Madagascar                        29 non-null     object
...
 52  Uganda                           29 non-null     object
 53  Venezuela                        29 non-null     object
 54  Viet Nam                         29 non-null     object
 55  Yemen                            29 non-null     object
dtypes: object(56)
memory usage: 12.9+ KB
```

PROJECT DATA: OVERVIEW

Prices Paid To Growers

```
prices_paid_to_growers.head()
```

prices_paid_to_growers	Colombia	Dominican Republic	El Salvador	Guatemala	Honduras	India	Uganda	Brazil	Ethiopia	India	Togo	Uganda	Other Nations
1990	1.534724	1.458168	1.116194	1.204956	1.11147	1.473558	0.337598	1.199223	1.348565	0.978921	0.645267	0.166486	0.943624
1991	1.48179	1.382845	0.983322	1.270086	1.238947	1.358371	0.654322	0.97115	1.505322	0.897289	0.632307	0.26143	0.964325
1992	1.204656	1.027841	0.682322	0.888099	0.886057	1.191159	0.441397	0.997768	1.351128	0.877945	0.658494	0.197653	0.761219
1993	1.106477	1.172704	0.780397	0.914552	0.828746	1.278669	0.552298	1.167263	1.362442	0.975912	0.499857	0.259737	0.806986
1994	1.898327	2.478234	2.191177	1.662711	1.800576	1.73081	1.666651	2.52911	2.418234	1.246437	0.573784	0.919709	1.585565

```
prices_paid_to_growers.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 29 entries, 1990 to 2018
Data columns (total 13 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Colombia            29 non-null     object
 1   Dominican Republic  29 non-null     object
 2   El Salvador         29 non-null     object
 3   Guatemala           29 non-null     object
 4   Honduras            29 non-null     object
 5   India               29 non-null     object
 6   Uganda              29 non-null     object
 7   Brazil              29 non-null     object
 8   Ethiopia            29 non-null     object
 9   India               29 non-null     object
 10  Togo                29 non-null     object
 11  Uganda              29 non-null     object
 12  Other Nations       29 non-null     float64
dtypes: float64(1), object(12)
memory usage: 3.2+ KB
```

ASSIGNMENT: MID-COURSE PROJECT

NEW MESSAGE
September 2024 ,18

From: **Clarissa Café** *(Coffee Client)*

Subject: **Summary Report**

Hi there,

Sarah told me to reach out directly to you – we loved the work you did on breaking down the industry, but we want to summarize your findings on Brazil into a single figure we can pass around.

Can you combine your findings into a single figure report? We'll also want to modify colors. There are more details in the attached notebook.

Thanks!
Clarissa

section05_coffee_project_part2.ipynb

Reply Forward

Key Objectives

1. Read in data from multiple csv files
2. Reshape the data with Pandas to set up charts
3. Build and customize line charts, bar charts, histograms and more to communicate key insights to our client
4. Modify chart colors to represent national flags
5. Combine modified charts into a single report by leveraging meshgrid and subplots

MCG

PROJECT 3: ADVANCED EDA

PROJECT DATA: USED CARS DATA

```
cars.head()
```

	year	make	model	trim	body	transmission	vin	state	condition	odometer	color	interior	seller	mmr	sellingprice	saledate
0	2015	Kia	Sorento	LX	SUV	automatic	5xyktca69fg566472	ca	5.0	16639.0	white	black	kia motors america, inc	20500	21500	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)
1	2015	Kia	Sorento	LX	SUV	automatic	5xyktca69fg561319	ca	5.0	9393.0	white	beige	kia motors america, inc	20800	21500	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)
2	2014	BMW	3 Series	328i SULEV	Sedan	automatic	wba3c1c51ek116351	ca	4.5	1331.0	gray	black	financial services remarketing (lease)	31900	30000	Thu Jan 15 2015 04:30:00 GMT-0800 (PST)
3	2015	Volvo	S60	T5	Sedan	automatic	yv1612tb4f1310987	ca	4.1	14282.0	white	black	volvo na rep/world omni	27500	27750	Thu Jan 29 2015 04:30:00 GMT-0800 (PST)
4	2014	BMW	6 Series Gran Coupe	650i	Sedan	automatic	wba6b2c57ed129731	ca	4.3	2641.0	gray	black	financial services remarketing (lease)	66000	67000	Thu Dec 18 2014 12:30:00 GMT-0800 (PST)

```
cars.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 558811 entries, 0 to 558810
Data columns (total 16 columns):
 #   Column        Non-Null Count   Dtype  
---  ------        --------------   -----  
 0   year          558811 non-null  int64  
 1   make          548510 non-null  object 
 2   model         548412 non-null  object 
 3   trim          548160 non-null  object 
 4   body          545616 non-null  object 
 5   transmission  493458 non-null  object 
 6   vin           558811 non-null  object 
 7   state         558811 non-null  object 
 8   condition     547017 non-null  float64
 9   odometer      558717 non-null  float64
 10  color         558062 non-null  object 
 11  interior      558062 non-null  object 
 12  seller        558811 non-null  object 
 13  mmr           558811 non-null  int64  
 14  sellingprice  558811 non-null  int64  
 15  saledate      558811 non-null  object 
dtypes: float64(2), int64(3), object(11)
memory usage: 68.2+ MB
```

ASSIGNMENT: FINAL PROJECT

NEW MESSAGE

October 2024 ,10

From: **Aaron Auto** *(VP of Fleet Management)*

Subject: **Optimal Fleet Truck Purchase**

Hello,

We need an outside analysis on auto procurement for our fleet of service vehicles. We lease trucks to contractors and other businesses, but a recent spike in demand has meant we're unable to get cars from traditional suppliers.

I want to see an overview of the automotive auction industry, before diving into where we can get Ford F150s for the most affordable price on the market (more details in the notebook).

Thanks

section07_final_project.ipynb

Reply Forward

Key Objectives

1. Read in and manipulate data with Pandas
2. Build summary charts with Matplotlib and Seaborn
3. Leverage Seaborn's advanced chart types to mine insights from the data and make a decision

MCG