```
In [1]:  import numpy as np
         import pandas as pd
```

```
In [2]:  df=pd.read_csv("hotel_booking_data.csv")
```

```
In [3]:  df
```

Out[3]:

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_nights	stays_in_week_nights	adults	...	customer_type	adr	requi
0	Resort Hotel	0	342	2015	July	27	1	0	0	2	...	Transient	0.00	
1	Resort Hotel	0	737	2015	July	27	1	0	0	2	...	Transient	0.00	
2	Resort Hotel	0	7	2015	July	27	1	0	1	1	...	Transient	75.00	
3	Resort Hotel	0	13	2015	July	27	1	0	1	1	...	Transient	75.00	
4	Resort Hotel	0	14	2015	July	27	1	0	2	2	...	Transient	98.00	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
119385	City Hotel	0	23	2017	August	35	30	2	5	2	...	Transient	96.14	
119386	City Hotel	0	102	2017	August	35	31	2	5	3	...	Transient	225.43	
119387	City Hotel	0	34	2017	August	35	31	2	5	2	...	Transient	157.71	
119388	City Hotel	0	109	2017	August	35	31	2	5	2	...	Transient	104.40	
119389	City Hotel	0	205	2017	August	35	29	2	7	2	...	Transient	151.20	

119390 rows × 36 columns

```
In [4]:  df.head()
```

Out[4]:

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_nights	stays_in_week_nights	adults	...	customer_type	adr	required_car_
0	Resort Hotel	0	342	2015	July	27	1	0	0	2	...	Transient	0.0	
1	Resort Hotel	0	737	2015	July	27	1	0	0	2	...	Transient	0.0	
2	Resort Hotel	0	7	2015	July	27	1	0	1	1	...	Transient	75.0	
3	Resort Hotel	0	13	2015	July	27	1	0	1	1	...	Transient	75.0	
4	Resort Hotel	0	14	2015	July	27	1	0	2	2	...	Transient	98.0	

5 rows × 36 columns

```
In [5]:  df.info()
```
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 119390 entries, 0 to 119389
Data columns (total 36 columns):
 #   Column                          Non-Null Count   Dtype
---  ------                          --------------   -----
 0   hotel                           119390 non-null  object
 1   is_canceled                     119390 non-null  int64
 2   lead_time                       119390 non-null  int64
 3   arrival_date_year               119390 non-null  int64
 4   arrival_date_month              119390 non-null  object
 5   arrival_date_week_number        119390 non-null  int64
 6   arrival_date_day_of_month       119390 non-null  int64
 7   stays_in_weekend_nights         119390 non-null  int64
 8   stays_in_week_nights            119390 non-null  int64
 9   adults                          119390 non-null  int64
 10  children                        119386 non-null  float64
 11  babies                          119390 non-null  int64
 12  meal                            119390 non-null  object
 13  country                         118902 non-null  object
 14  market_segment                  119390 non-null  object
 15  distribution_channel            119390 non-null  object
 16  is_repeated_guest               119390 non-null  int64
 17  previous_cancellations          119390 non-null  int64
 18  previous_bookings_not_canceled  119390 non-null  int64
 19  reserved_room_type              119390 non-null  object
 20  assigned_room_type              119390 non-null  object
 21  booking_changes                 119390 non-null  int64
 22  deposit_type                    119390 non-null  object
 23  agent                           103050 non-null  float64
 24  company                         6797 non-null    float64
 25  days_in_waiting_list            119390 non-null  int64
 26  customer_type                   119390 non-null  object
 27  adr                             119390 non-null  float64
 28  required_car_parking_spaces     119390 non-null  int64
 29  total_of_special_requests       119390 non-null  int64
 30  reservation_status              119390 non-null  object
 31  reservation_status_date         119390 non-null  object
 32  name                            119390 non-null  object
 33  email                           119390 non-null  object
 34  phone-number                    119390 non-null  object
 35  credit_card                     119390 non-null  object
dtypes: float64(4), int64(16), object(16)
memory usage: 32.8+ MB
```

1.What are the top 5 most common country codesin the dataset?

```
In [6]:  df["country"].value_counts()[:5]
```
```
Out[6]:  PRT    48590
         GBR    12129
         FRA    10415
         ESP     8568
         DEU     7287
         Name: country, dtype: int64
```

3.what is the average(mean)number of nights of a stay across the entire data set?

```
In [11]:  df['total_stay_days']=df['stays_in_week_nights']+df['stays_in_weekend_nights']
          df['total_stay_days'].mean()
```
```
Out[11]:  3.4279001591423066
```

4.what is the avreage total cost for a stay in the dataset?

```
In [12]:  df['total_paid']=df['adr']*df['total_stay_days']
          df['total_paid'].mean()
```
```
Out[12]:  357.8482078063524
```

5.what are the names and email of people who made 5"special requests"?

```
In [14]:  df[df['total_of_special_requests']==5][['name','email']]
```

Out[14]:

	name	email
7860	Amanda Harper	Amanda.H66@yahoo.com
11125	Laura Sanders	Sanders_Laura@hotmail.com
14596	Tommy Ortiz	Tommy_O@hotmail.com
14921	Gilbert Miller	Miller.Gilbert@aol.com
14922	Timothy Torres	TTorres@protonmail.com
24630	Jennifer Weaver	Jennifer_W@aol.com
27288	Crystal Horton	Crystal.H@mail.com
27477	Brittney Burke	Burke_Brittney16@att.com
29906	Cynthia Cabrera	Cabrera.Cynthia@xfinity.com
29949	Sarah Floyd	Sarah_F@gmail.com
32267	Michelle Villa	Michelle.Villa@aol.com
39027	Nichole Hebert	Hebert.Nichole@gmail.com
39129	Lindsey Mckenzie	Lindsey.Mckenzie@att.com
39525	Ashley Edwards	Edwards.Ashley@yahoo.com
70114	Christopher Torres	Torres.Christopher@gmail.com
78819	Mrs. Tara Sullivan DVM	Mrs._DVM@xfinity.com
78820	Michaela Brown	MichaelaBrown@att.com
78822	Kurt Maldonado MD	KMD15@xfinity.com
97072	Jason Richardson	Jason.R@zoho.com
97099	Terri Hurley	THurley@xfinity.com
97261	Mrs. Caitlin Webb	Mrs._W@comcast.net
98410	Holly Arroyo	Arroyo_Holly@mail.com
98674	Denise Campbell	Denise_C@gmail.com
99887	Michael Smith	Michael.S42@aol.com
99888	Dr. Trevor Sellers	Dr._S@aol.com
101569	Kayla Murphy	Kayla.Murphy@yahoo.com
102061	Taylor Martinez	Taylor.Martinez@hotmail.com
109511	Charles Wilson	Charles_Wilson@yahoo.com
109590	Tyler Allison	Tyler.A@protonmail.com
110082	Matthew Bailey	Matthew_Bailey@aol.com
110083	Charlotte Acevedo	Charlotte_A@verizon.com
111909	Darrell Brennan	Brennan_Darrell51@hotmail.com
111911	Melinda Jensen	MelindaJensen@zoho.com
113915	Terry Arnold	Arnold.Terry@zoho.com
114770	Mary Nguyen	Nguyen.Mary@protonmail.com
114909	Lindsay Cuevas	Lindsay.Cuevas40@mail.com
116455	Cynthia Hernandez	CynthiaHernandez@xfinity.com
116457	Angela Hawkins	Angela_H@gmail.com
118817	Sue Lawson	Sue.L52@comcast.net
119161	Alyssa Richards	Alyssa_Richards@aol.com

2.what is the name of the person who paid the highest ADR(average daily rate)?How much is their ADR?

```
In [18]:  df.sort_values('adr',ascending=False)[['adr','name']].iloc[0]
```
```
Out[18]:  adr            5400.0
          name    Daniel Walter
          Name: 48515, dtype: object
```

```
In [ ]:
```