In [8]:
```python
import pandas as pd
```

In [9]:
```python
data = pd.read_csv("TWO_CENTURIES_OF_UM_RACES.csv")
```

```
C:\Users\Manisha\AppData\Local\Temp\ipykernel_18804\1957940255.py:1: DtypeWarning: Columns (11) have mixed types. Specify dtype option on import or set low_memory=False.
  data = pd.read_csv("TWO_CENTURIES_OF_UM_RACES.csv")
```

In [10]:
```python
# Show 10 rows
```

In [11]:
```python
data.head(10)
```

Out[11]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete club	Athlete country	Athlete year of birth	Athlete gender	Athlete age category	Athlete average speed	Athlete ID
0	2018	06.01.2018	Selva Costera (CHI)	50km	22	4:51:39 h	Tnfrc	CHI	1978.0	M	M35	10.286	0
1	2018	06.01.2018	Selva Costera (CHI)	50km	22	5:15:45 h	Roberto Echeverría	CHI	1981.0	M	M35	9.501	1
2	2018	06.01.2018	Selva Costera (CHI)	50km	22	5:16:44 h	Puro Trail Osorno	CHI	1987.0	M	M23	9.472	2
3	2018	06.01.2018	Selva Costera (CHI)	50km	22	5:34:13 h	Columbia	ARG	1976.0	M	M40	8.976	3
4	2018	06.01.2018	Selva Costera (CHI)	50km	22	5:54:14 h	Baguales Trail	CHI	1992.0	M	M23	8.469	4
5	2018	06.01.2018	Selva Costera (CHI)	50km	22	6:25:01 h	NaN	ARG	1974.0	M	M40	7.792	5
6	2018	06.01.2018	Selva Costera (CHI)	50km	22	6:28:00 h	Los Patagones	ARG	1979.0	F	W35	7.732	6
7	2018	06.01.2018	Selva Costera (CHI)	50km	22	6:32:24 h	Reaktiva Chile	CHI	1967.0	F	W50	7.645	7
8	2018	06.01.2018	Selva Costera (CHI)	50km	22	6:39:08 h	Puro Trail Osorno	CHI	1985.0	M	M23	7.516	8
9	2018	06.01.2018	Selva Costera (CHI)	50km	22	6:45:11 h	Marlene Flores Team	CHI	1976.0	M	M40	7.404	9

In [12]:
```python
data.shape
```

Out[12]:
```
(7461195, 13)
```

In [13]:
```python
data.dtypes
```

Out[13]:
```
Year of event                  int64
Event dates                   object
Event name                    object
Event distance/length         object
Event number of finishers      int64
Athlete performance           object
Athlete club                  object
Athlete country               object
Athlete year of birth        float64
Athlete gender                object
Athlete age category          object
Athlete average speed         object
Athlete ID                     int64
dtype: object
```

In [14]:
```python
# show 60km event distance
```

In [15]:
```python
data[data["Event distance/length"]=="60km"]
```

Out[15]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete club	Athlete country	Athlete year of birth	Athlete gender	Athlete age category	Athlete average speed	Athlete ID
1868	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	7:01:49 h	NaN	UAE	1972.0	M	M45	8.535	1754
1869	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	7:10:09 h	NaN	FRA	1985.0	F	W23	8.369	1755
1870	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	7:39:41 h	NaN	GBR	1971.0	M	M45	7.831	1756
1871	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	8:21:36 h	NaN	GBR	1978.0	F	W35	7.177	1757
1872	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	8:36:08 h	NaN	UAE	1985.0	M	M23	6.975	1758
...	...	...	...	...	...	...	...	...	...	...	...	...	...
7461030	1995	08.01.1995	Bogong to Hotham (AUS)	60km	11	9:51:28 h	NaN	AUS	1958.0	M	M35	6087.0	1062126
7461031	1995	08.01.1995	Bogong to Hotham (AUS)	60km	11	9:51:28 h	NaN	AUS	1945.0	M	M45	6087.0	1046173
7461032	1995	08.01.1995	Bogong to Hotham (AUS)	60km	11	9:51:28 h	NaN	AUS	1955.0	M	M35	6087.0	1046178
7461033	1995	08.01.1995	Bogong to Hotham (AUS)	60km	11	9:51:28 h	NaN	AUS	1966.0	M	M23	6087.0	1044961
7461034	1995	08.01.1995	Bogong to Hotham (AUS)	60km	11	10:24:13 h	NaN	AUS	1930.0	M	M60	5767.0	1070120

159494 rows × 13 columns

In [16]:
```python
# show female athletes who runned for 60 km
```

In [17]:
```python
data[(data["Event distance/length"]=="60km") & (data["Athlete gender"] == "F")]
```

Out[17]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete club	Athlete country	Athlete year of birth	Athlete gender	Athlete age category	Athlete average speed	Athlete ID
1869	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	7:10:09 h	NaN	FRA	1985.0	F	W23	8.369	1755
1871	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	8:21:36 h	NaN	GBR	1978.0	F	W35	7.177	1757
1875	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	9:31:32 h	NaN	USA	1974.0	F	W40	6.299	1761
1879	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	9:56:47 h	NaN	GBR	1990.0	F	W23	6.032	1765
1881	2018	23.03.2018	Urban-Ultra Mt. Sana 60 (UAE)	60km	24	10:22:30 h	NaN	USA	1981.0	F	W35	5.783	1767
...	...	...	...	...	...	...	...	...	...	...	...	...	...
7460335	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	6:32:17 h	*New York, NY	USA	1954.0	F	W40	9177.0	1050214
7460336	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	6:42:31 h	*Bronx, NY	USA	1969.0	F	W23	8944.0	1641118
7460337	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	6:42:55 h	*New York, NY	USA	1967.0	F	W23	8935.0	1050222
7460343	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	8:08:33 h	RR	USA	1945.0	F	W45	7369.0	1641119
7461027	1995	08.01.1995	Bogong to Hotham (AUS)	60km	11	8:29:38 h	NaN	AUS	1965.0	F	W23	7064.0	1044891

31565 rows × 13 columns

In [18]:
```python
# Show event name for NYRR Knickerbocker 60K (USA)
```

In [19]:
```python
data[data["Event name"]== "NYRR Knickerbocker 60K (USA)"]
```

Out[19]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete club	Athlete country	Athlete year of birth	Athlete gender	Athlete age category	Athlete average speed	Athlete ID
624622	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:24:36 h	NaN	PAR	1970.0	M	M45	13.605	359326
624623	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:26:17 h	*New York, NY	USA	1984.0	M	M23	13.519	359327
624624	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:37:47 h	*Scotch Plains, NJ	USA	1983.0	M	M23	12.96	46880
624625	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:40:20 h	*New York, NY	USA	1991.0	M	M23	12.842	359328
624626	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:44:39 h	*Brooklyn, NY	USA	1982.0	M	M35	12.647	359329
...	...	...	...	...	...	...	...	...	...	...	...	...	...
7460339	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	7:02:23 h	*New York, NY	USA	1935.0	M	M55	8523.0	1050218
7460340	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	7:10:44 h	*CT	USA	1951.0	M	M40	8358.0	1194267
7460341	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	7:22:41 h	PPTC	USA	1943.0	M	M50	8132.0	1046034
7460342	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	8:02:47 h	APS	USA	1916.0	M	M75	7457.0	1069921
7460343	1995	25.02.1995	NYRR Knickerbocker 60K (USA)	60km	43	8:08:33 h	RR	USA	1945.0	F	W45	7369.0	1641119

4936 rows × 13 columns

In [20]:
```python
# describe the data
```

In [21]:
```python
data[data["Event name"]== "NYRR Knickerbocker 60K (USA)"].describe()
```

Out[21]:

	Year of event	Event number of finishers	Athlete year of birth	Athlete ID
count	4936.000000	4936.000000	4930.000000	4.936000e+03
mean	2009.251621	243.480956	1969.149493	7.336843e+05
std	9.411665	123.322610	14.263588	4.907308e+05
min	1980.000000	33.000000	1916.000000	8.600000e+01
25%	2006.000000	98.000000	1959.250000	3.593718e+05
50%	2012.000000	297.000000	1971.000000	6.381625e+05
75%	2016.000000	358.000000	1980.000000	1.196144e+06
max	2019.000000	375.000000	1999.000000	1.641119e+06

In [22]:
```python
data[data["Event name"]== "NYRR Knickerbocker 60K (USA)"]["Event name"].str.split("(").str.get(1).str.split("(")
```

Out[22]:
```
624622     [USA)]
624623     [USA)]
624624     [USA)]
624625     [USA)]
624626     [USA)]
            ...
7460339    [USA)]
7460340    [USA)]
7460341    [USA)]
7460342    [USA)]
7460343    [USA)]
Name: Event name, Length: 4936, dtype: object
```

In [23]:
```python
data[(data["Event name"]=="NYRR Knickerbocker 60K (USA)") & (data["Year of event"]== 2018)]
```

Out[23]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete club	Athlete country	Athlete year of birth	Athlete gender	Athlete age category	Athlete average speed	Athlete ID
624622	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:24:36 h	NaN	PAR	1970.0	M	M45	13.605	359326
624623	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:26:17 h	*New York, NY	USA	1984.0	M	M23	13.519	359327
624624	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:37:47 h	*Scotch Plains, NJ	USA	1983.0	M	M23	12.96	46880
624625	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:40:20 h	*New York, NY	USA	1991.0	M	M23	12.842	359328
624626	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:44:39 h	*Brooklyn, NY	USA	1982.0	M	M35	12.647	359329
...	...	...	...	...	...	...	...	...	...	...	...	...	...
624977	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	9:52:01 h	*Metuchen, NJ	USA	1976.0	F	W40	6.081	359576
624978	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	9:52:03 h	*Ozone Park, NY	USA	1969.0	F	W45	6.081	359577
624979	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	10:06:56 h	NaN	IND	1984.0	F	W23	5.931	359578
624980	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	10:07:08 h	*New York, NY	USA	1960.0	M	M55	5.93	359579
624981	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	10:26:58 h	*New York, NY	USA	1942.0	F	W75	5.742	359580

360 rows × 13 columns

In [24]:
```python
# Combine all filters together
```

In [25]:
```python
df= data[(data["Event name"]=="NYRR Knickerbocker 60K (USA)") & (data["Year of event"]== 2018)]
```

In [26]:
```python
df.head(8)
```

Out[26]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete club	Athlete country	Athlete year of birth	Athlete gender	Athlete age category	Athlete average speed	Athlete ID
624622	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:24:36 h	NaN	PAR	1970.0	M	M45	13.605	359326
624623	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:26:17 h	*New York, NY	USA	1984.0	M	M23	13.519	359327
624624	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:37:47 h	*Scotch Plains, NJ	USA	1983.0	M	M23	12.96	46880
624625	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:40:20 h	*New York, NY	USA	1991.0	M	M23	12.842	359328
624626	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:44:39 h	*Brooklyn, NY	USA	1982.0	M	M35	12.647	359329
624627	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:45:12 h	*Jackson Heights, NY	USA	1974.0	M	M40	12.623	359330
624628	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:46:08 h	*Summit, NJ	USA	1973.0	M	M45	12.582	40939
624629	2018	17.11.2018	NYRR Knickerbocker 60K (USA)	60km	360	4:48:05 h	*New York, NY	USA	1986.0	M	M23	12.496	359331

In [27]:
```python
df.shape
```

Out[27]:
```
(360, 13)
```

In [28]:
```python
# remove USA from event name
```

In [29]:
```python
df["Event name"].str.split("(").str.get(0)
```

Out[29]:
```
624622    NYRR Knickerbocker 60K 
624623    NYRR Knickerbocker 60K 
624624    NYRR Knickerbocker 60K 
624625    NYRR Knickerbocker 60K 
624626    NYRR Knickerbocker 60K 
                   ...
624977    NYRR Knickerbocker 60K 
624978    NYRR Knickerbocker 60K 
624979    NYRR Knickerbocker 60K 
624980    NYRR Knickerbocker 60K 
624981    NYRR Knickerbocker 60K 
Name: Event name, Length: 360, dtype: object
```

In [30]:
```python
df["Event name"] = df["Event name"].str.split("(").str.get(0)
```

```
C:\Users\Manisha\AppData\Local\Temp\ipykernel_18804\2314686612.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df["Event name"] = df["Event name"].str.split("(").str.get(0)
```

In [31]:
```python
df.head(8)
```

Out[31]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete club	Athlete country	Athlete year of birth	Athlete gender	Athlete age category	Athlete average speed	Athlete ID
624622	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:24:36 h	NaN	PAR	1970.0	M	M45	13.605	359326
624623	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:26:17 h	*New York, NY	USA	1984.0	M	M23	13.519	359327
624624	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:37:47 h	*Scotch Plains, NJ	USA	1983.0	M	M23	12.96	46880
624625	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:40:20 h	*New York, NY	USA	1991.0	M	M23	12.842	359328
624626	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:44:39 h	*Brooklyn, NY	USA	1982.0	M	M35	12.647	359329
624627	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:45:12 h	*Jackson Heights, NY	USA	1974.0	M	M40	12.623	359330
624628	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:46:08 h	*Summit, NJ	USA	1973.0	M	M45	12.582	40939
624629	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:48:05 h	*New York, NY	USA	1986.0	M	M23	12.496	359331

In [32]:
```python
# remove Athlete club,Athlete country,Athlete year of birth,Athlete gender,Athlete age category
```

In [33]:
```python
df = df.drop(["Athlete club","Athlete country","Athlete year of birth","Athlete gender","Athlete age category"], axis =1)
```

In [34]:
```python
df.head(5)
```

Out[34]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete average speed	Athlete ID
624622	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:24:36 h	13.605	359326
624623	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:26:17 h	13.519	359327
624624	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:37:47 h	12.96	46880
624625	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:40:20 h	12.842	359328
624626	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:44:39 h	12.647	359329

In [35]:
```python
# check duplicate values
```

In [36]:
```python
df[df.duplicated()== True]  # their is no duplicate values
```

Out[36]:

Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete average speed	Athlete ID

In [ ]:

In [37]:
```python
# reset index
```

In [38]:
```python
df.reset_index(drop = True)
```

Out[38]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete average speed	Athlete ID
0	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:24:36 h	13.605	359326
1	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:26:17 h	13.519	359327
2	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:37:47 h	12.96	46880
3	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:40:20 h	12.842	359328
4	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:44:39 h	12.647	359329
...	...	...	...	...	...	...	...	...
355	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	9:52:01 h	6.081	359576
356	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	9:52:03 h	6.081	359577
357	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	10:06:56 h	5.931	359578
358	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	10:07:08 h	5.93	359579
359	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	10:26:58 h	5.742	359580

360 rows × 8 columns

In [39]:
```python
df.dtypes
```

Out[39]:
```
Year of event                  int64
Event dates                   object
Event name                    object
Event distance/length         object
Event number of finishers      int64
Athlete performance           object
Athlete average speed         object
Athlete ID                     int64
dtype: object
```

In [40]:
```python
# change types
```

In [41]:
```python
df["Athlete ID"]= df["Athlete ID"].astype("float")
```

In [42]:
```python
df.dtypes
```

Out[42]:
```
Year of event                  int64
Event dates                   object
Event name                    object
Event distance/length         object
Event number of finishers      int64
Athlete performance           object
Athlete average speed         object
Athlete ID                   float64
dtype: object
```

In [43]:
```python
df.head(5)
```

Out[43]:

	Year of event	Event dates	Event name	Event distance/length	Event number of finishers	Athlete performance	Athlete average speed	Athlete ID
624622	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:24:36 h	13.605	359326.0
624623	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:26:17 h	13.519	359327.0
624624	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:37:47 h	12.96	46880.0
624625	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:40:20 h	12.842	359328.0
624626	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:44:39 h	12.647	359329.0

In [44]:
```python
# rename colums : Year of event, Event dates , Event name, Event number
```

In [45]:
```python
df=df.rename(columns={"Year of event": "Year event",
                      "Event Dates": "event dates",
                      "Event name":  "event name",
                      "Event number of finishers" : "event number of finish"})
```

In [46]:
```python
df.head(4)
```

Out[46]:

	Year event	Event dates	event name	Event distance/length	event number of finish	Athlete performance	Athlete average speed	Athlete ID
624622	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:24:36 h	13.605	359326.0
624623	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:26:17 h	13.519	359327.0
624624	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:37:47 h	12.96	46880.0
624625	2018	17.11.2018	NYRR Knickerbocker 60K	60km	360	4:40:20 h	12.842	359328.0

In [47]:
```python
# Reorder columns
```

In [53]:
```python
df2=df[["Event dates", "Year event","Event distance/length", "event name","event number of finish", "Athlete ID","Athlete average speed","Athlete pe
```

In [55]:
```python
df2.head(10)
```

Out[55]:

	Event dates	Year event	Event distance/length	event name	event number of finish	Athlete ID	Athlete average speed	Athlete performance
624622	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	359326.0	13.605	4:24:36 h
624623	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	359327.0	13.519	4:26:17 h
624624	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	46880.0	12.96	4:37:47 h
624625	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	359328.0	12.842	4:40:20 h
624626	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	359329.0	12.647	4:44:39 h
624627	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	359330.0	12.623	4:45:12 h
624628	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	40939.0	12.582	4:46:08 h
624629	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	359331.0	12.496	4:48:05 h
624630	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	298539.0	12.38	4:50:48 h
624631	17.11.2018	2018	60km	NYRR Knickerbocker 60K	360	29389.0	12.375	4:50:55 h

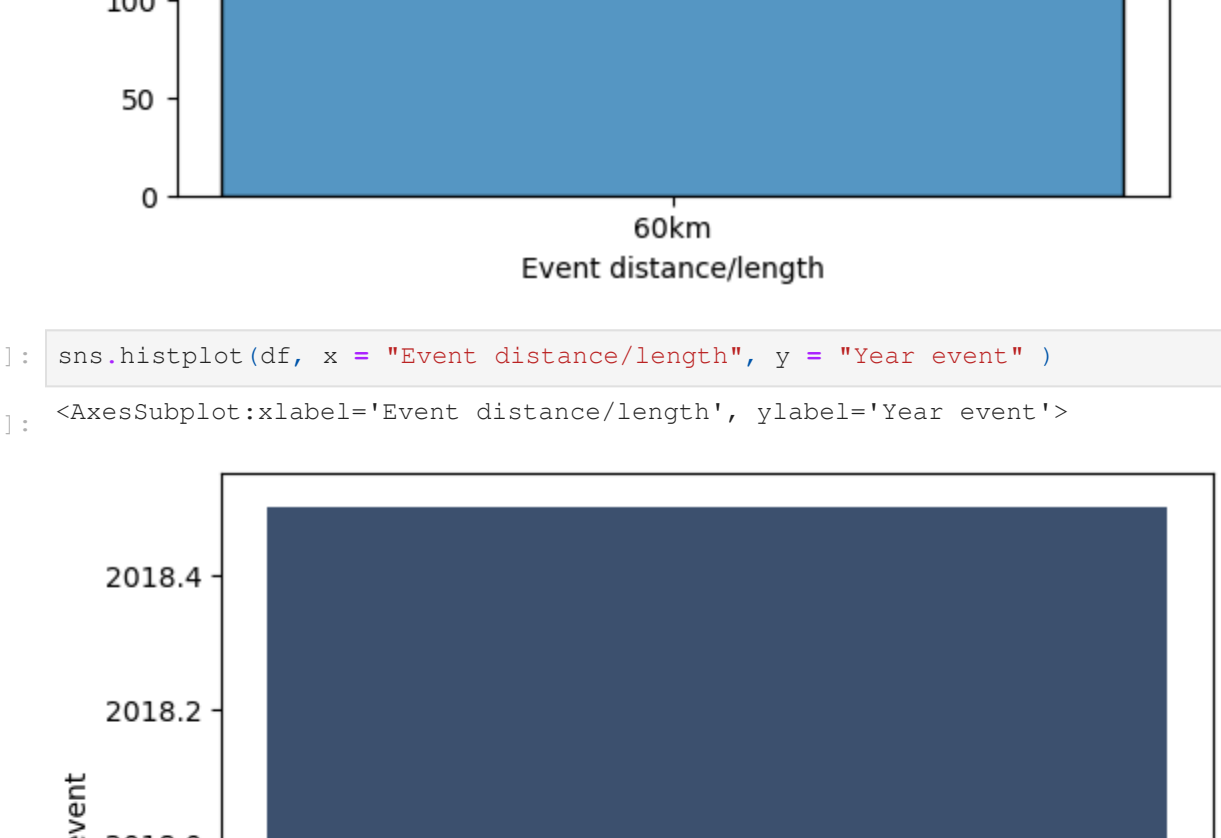
In [ ]:
```python
# show event name with year 2018 # 359329.0
```

In [58]:
```python
import seaborn as sns
```

In [ ]:
```python
# charts and graph
```

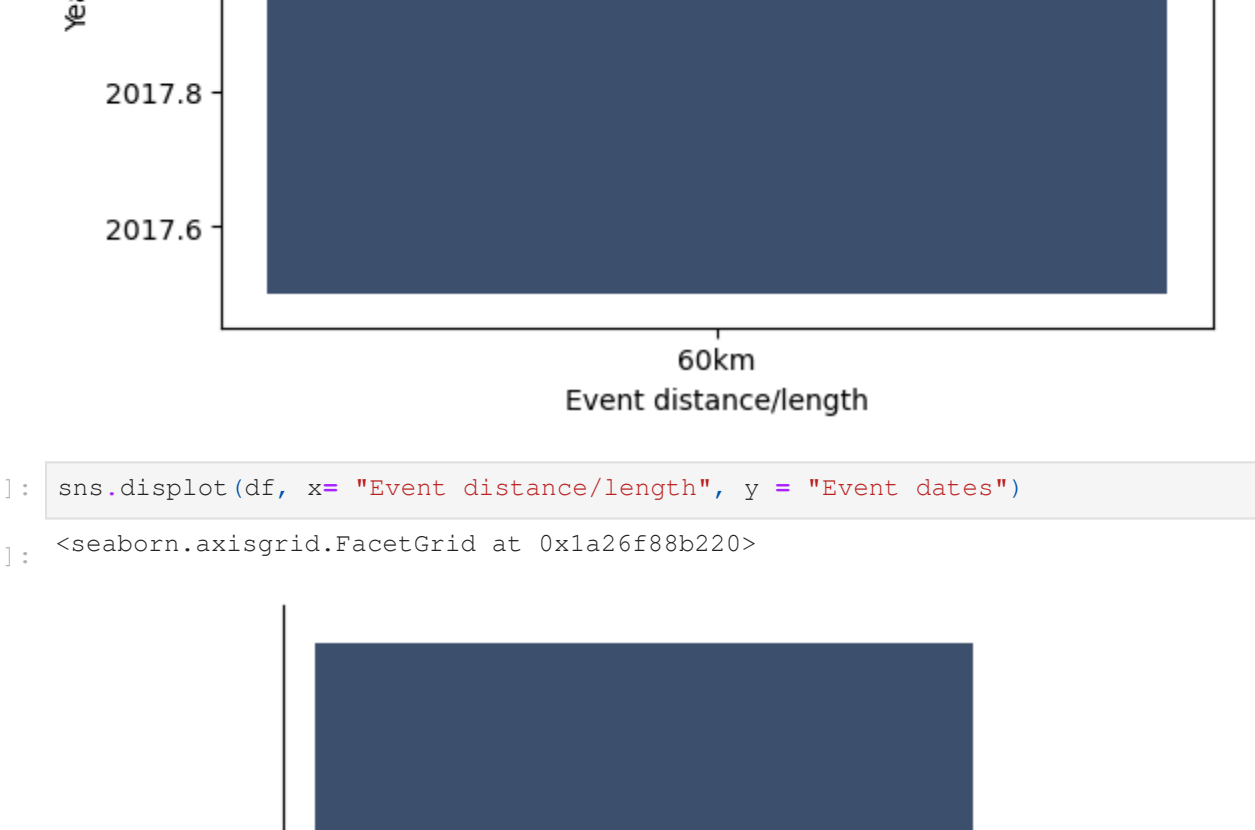
In [59]:
```python
sns.histplot(df["Event distance/length"])
```

Out[59]:
```
<AxesSubplot:xlabel='Event distance/length', ylabel='Count'>
```



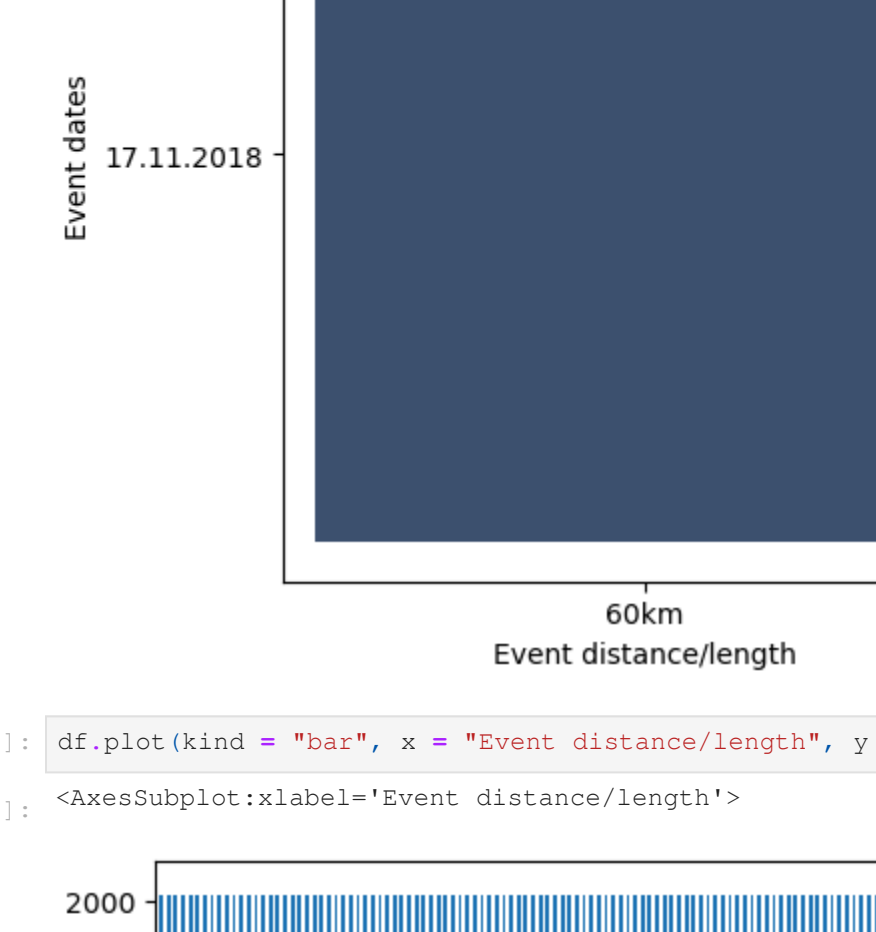
In [63]:
```python
sns.histplot(df, x = "Event distance/length", y = "Year event" )
```

Out[63]:
```
<AxesSubplot:xlabel='Event distance/length', ylabel='Year event'>
```



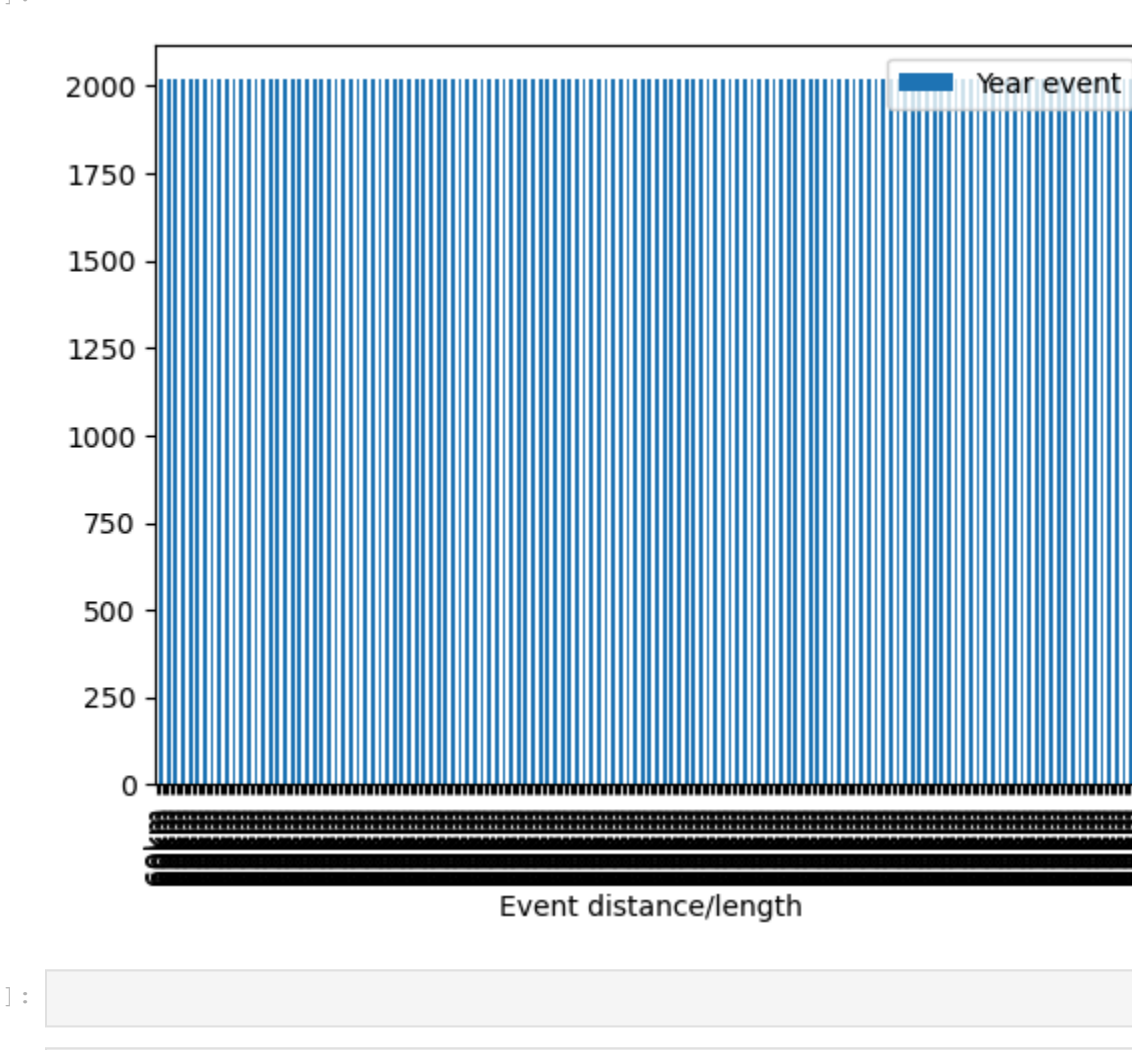
In [66]:
```python
sns.displot(df, x= "Event distance/length", y = "Event dates")
```

Out[66]:
```
<seaborn.axisgrid.FacetGrid at 0x1a26f88b220>
```



In [74]:
```python
df.plot(kind = "bar", x = "Event distance/length", y = "Year event")
```

Out[74]:
```
<AxesSubplot:xlabel='Event distance/length'>
```



In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: