

## Avistamiento de Ovnis

### ❖ Introducción.

Desde hace mucho tiempo el hombre se ha preguntado sobre la existencia de la vida en otros planetas, muchas personas dedican su vida entera a la búsqueda de señales o avistamientos que puedan ayudarlos a confirmar dicha existencia.

En esta ocasión nos toca colaborar en la determinación del lugar con mayor probabilidad que se pueda presentar algún avistamiento dado los reportes.

Contamos con una gran cantidad de datos que nos indica, la fecha, el lugar preciso, forma y duración del avistamiento, como se muestra a continuación.

	datetime	city	state	country	shape	duration (seconds)	duration (hours/min)	comments	date posted	latitude	longitude
0	10/10/1949 20:30	san marcos	tx	us	cylinder	2700	45 minutes	This event took place in early fall around 194...	4/27/2004	29.8830556	-97.941111
1	10/10/1949 21:00	lackland afb	tx	NaN	light	7200	1-2 hrs	1949 Lackland AFB&#44 TX Lights racing across...	12/16/2005	29.38421	-98.581082
2	10/10/1955 17:00	chester (uk/england)	NaN	gb	circle	20	20 seconds	Green/Orange circular disc over Chester&#44 En...	1/21/2008	53.2	-2.916667
3	10/10/1956 21:00	edna	tx	us	circle	20	1/2 hour	My older brother and twin sister were leaving ...	1/17/2004	28.9783333	-96.645833
4	10/10/1960 20:00	kaneohe	hi	us	light	900	15 minutes	AS a Marine 1st Lt. flying an FJ4B fighter/att...	1/22/2004	21.4180556	-157.803611

tabla 1. Primeras 4 filas de data set

### ❖ Exploración y limpieza de datos.

En la exploración de los datos se encontraron varias inconsistencias.

- En la columna "latitude", fila "43782" se encontró un dato con formato erróneo debido a que estaba mal escrito, se reescribió de acuerdo a un rango válido de valores.
- En la columna "duration (seconds)" se encontraron 3 valores en las filas "27822, 35692 y 58591" con un formato erróneo, apoyado de la columna "duration" (hours/min) se reescribieron los valores para tener el formato correcto.
- Se asignó un formato de tipo "datetime" en la columna date "posted".

A continuación, se muestra el tipo de los datos por columna de acuerdo a la biblioteca de datos.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 80332 entries, 0 to 80331
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   datetime                             80332 non-null  object
1   city                                80332 non-null  object
2   state                               74535 non-null  object
3   country                             70662 non-null  object
4   shape                               78400 non-null  object
5   duration (seconds)                  80332 non-null  float64
6   duration (hours/min)                80332 non-null  object
7   comments                            80317 non-null  object
8   date posted                         80332 non-null  datetime64[ns]
9   latitude                           80332 non-null  float64
10  longitude                           80332 non-null  float64
dtypes: datetime64[ns](1), float64(3), object(7)
memory usage: 6.7+ MB

```

tabla 2. Tipo de datos y numero de datos por columna.

#### ❖ Análisis.

Una vez terminado la limpieza de datos, podemos visualizar de forma gráfica los lugares con mayor cantidad de avistamientos por país, estado y ciudad.

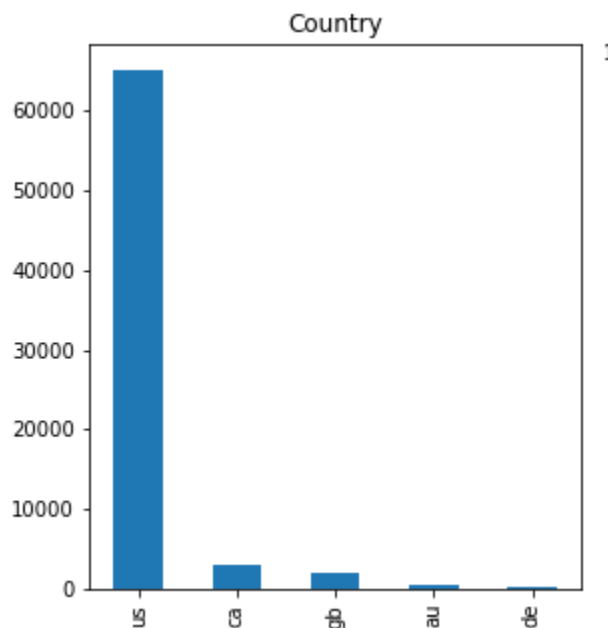


Grafico.1 Pasises con mayor cantidad de avistamientos.

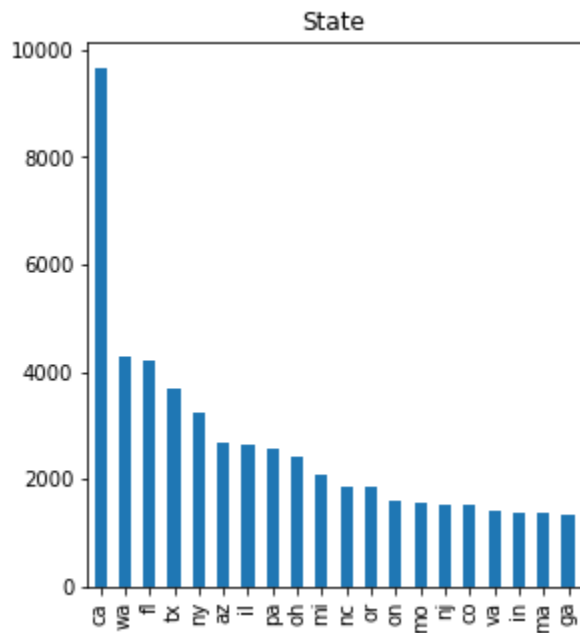


Gráfico.2 Estados con mayor cantidad de avistamientos.

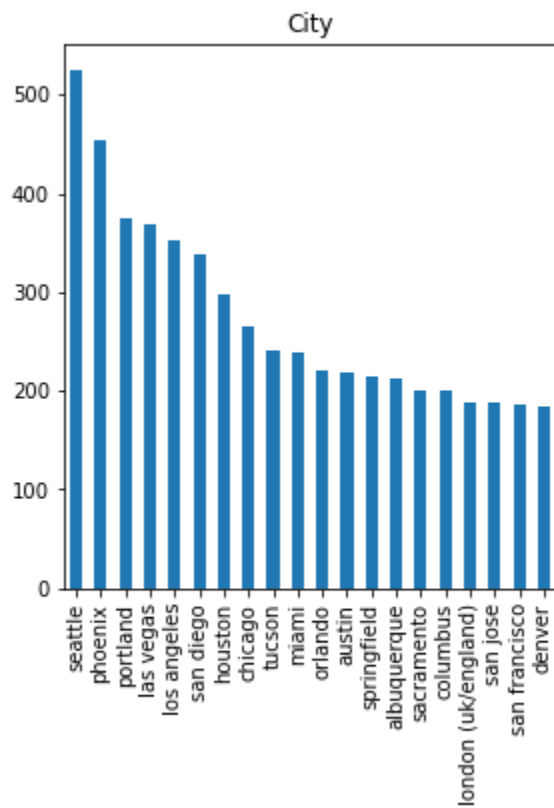


Grafico.3 Ciudades con mayor cantidad de avistamientos.

En el grafico 1 se observa a Estados Unidos “us” como el país como mayor número de avistamientos reportados.

En el grafico 2 se observa a California “ca” como el estado como mayor número de avistamientos reportados.

Por último del grafico 3 se observa a Seattle “seattle” como la ciudad con mayor número de avistamientos reportados.

A manera de comprobación se muestran la cantidad exacta de avistamientos reportados respectivamente.

País: us 65114  
Estado: ca 9655  
Ciudad: seattle 525

Para la toma de una decisión final se considero que tanto el estado como la ciudad pertenecen al mismo país.

Considerando el estado (California “ca”) como mayor lugar con avistamientos reportados, se determina la ciudad con mayor número, siendo esta (los ángeles)

los angeles	352
san diego	336
sacramento	201
san jose	186
san francisco	186
...	

Tabla 3. Ciudades de California (ca) con mayor número de avistamientos.

Notando que se encuentra dentro de las 5 primeras ciudades con mayor número de avistamientos (ver gráfico 3).

Considerando solo la ciudad se concluye directamente que es Seattle.

En comparación con California hay una diferencia de 173 avistamientos.

#### ❖ Resultado.

El lugar que cuenta con mayor probabilidad de presenciar un avistamiento es la ciudad de Seattle.