



Benjamín Rivera
**Evidencia de aprendizaje: Reporte
estadístico**

Universidad Abierta y a Distancia de México

TSU en Biotecnología

Materia: Estadística Básica

Grupo: BI-BEBA-2002-B2-013

Unidad: Unidad 3

Matricula: ES202105994

Fecha de entrega: 24 de noviembre de 2020

1. Caso de estudio

Estoy interesado en estudiar las **técnicas de repoblación de ecosistemas** usando agentes biológicos. Uno de los temas de interes para poder desarrollar esta investigación es conocer las **especies por espacio geográfico** que habitan. Esto es importante para poder analizar ecosistemas afectados y aquellos que sean sanos, y que posean características similares.



Figura 1: Corales muertos en el Océano Índico por culpa del calentamiento global [?]

Entonces en **esta parte** nos centraremos en las **ubicaciones de los bancos de algas del mundo**. Esta información nos pertmitira teorizar técnicas de *biorremediación* en función de otros ambientes similares. Además, el estudio de estos en algún tiempo determinado nos dara una oportunidad para identificar causas de infeccion y muerte de las agrupaciones de estos; para tratar de predecir los corales que esten en peligro por causas similares.

2. Base de Datos

De manera que usando la base de datos [3] obtenemos la visualización preliminar que podemos ver en la figura 2.

ID	latitude	longitude	Ocean	Realm	Ecoregion	Country_Name	State_Island_Prov	depth	Average_Bleach	ClimSST	Temperature_Kelvin	Temperature_M	Temperature_M	Temperature_M	Temperature_Kelvin
unlless	degrees_north	degrees_east	unlless	unlless	unlless	unlless	unlless	m	percent	Degrees Celsius	Kelvin	Degree Celsius	Degree Celsius	Degree Celsius	Kelvin
97	-20.89963333	149.4077222	Pacific	Central Indo-Pac Southern Great E Australia	Queensland			4	6.25	297.78	301.67	298.17	291.91	303.85	2.74
98	-20.89305556	149.421	Pacific	Central Indo-Pac Southern Great E Australia	Queensland			5	8.75	297.78	301.67	298.17	291.91	303.85	2.74
116	-20.74580556	149.4720556	Pacific	Central Indo-Pac Southern Great E Australia	Queensland			6	11.25	298.17	301.38	298.34	292.67	303.76	2.57
117	-20.73769444	149.4649444	Pacific	Central Indo-Pac Southern Great E Australia	Queensland			5	25	298.21	301.56	298.37	292.01	304.04	2.52
142	-20.25906111	148.8145833	Pacific	Central Indo-Pac Central and north Australia	Queensland			2	1	296.48	296.38	298.47	291.83	304.54	2.7
145	-20.25255556	148.8129167	Pacific	Central Indo-Pac Central and north Australia	Queensland			2	33.75	296.48	296.38	298.47	291.83	304.54	2.7
155	-20.06580556	148.9481111	Pacific	Central Indo-Pac Central and north Australia	Queensland			7	9.25	298.77	299.31	298.68	293.39	304.04	2.43
158	-20.039	148.8758333	Pacific	Central Indo-Pac Central and north Australia	Queensland			4	40	296.79	299.88	298.69	292.94	304.3	2.48
166	-19.81777778	149.0658056	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	9.5	295.74	299.34	298.72	292.92	304.14	2.23
167	-19.81777778	149.0658056	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	11.75	296.83	296.51	298.72	292.92	304.14	2.23
168	-19.81311111	149.0657222	Pacific	Central Indo-Pac Central and north Australia	Queensland			2	1.25	296.83	296.51	298.72	292.92	304.14	2.23
169	-19.80755556	149.0627222	Pacific	Central Indo-Pac Central and north Australia	Queensland			4	4.75	295.74	299.34	298.72	292.92	304.14	2.23
172	-19.74577778	149.1659444	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	32.5	298.51	298.76	298.59	293.64	303.52	2.24
173	-19.74577778	149.1659444	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	0	298.59	296.16	298.59	293.64	303.52	2.24
174	-19.74555556	149.1658333	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	17.25	298.51	298.76	298.59	293.64	303.52	2.24
175	-19.74555556	149.1658333	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	1.25	298.59	296.16	298.59	293.64	303.52	2.24
183	-19.29	-158.9583333	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Manuae			10	0	298.72	298.79	299.67	296.01	304.19	1.31
184	-19.26833333	-158.975	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Manuae			10	0	298.74	298.91	299.69	295.67	303.92	1.31
185	-19.19833333	146.8151667	Pacific	Central Indo-Pac Central and north Australia	Queensland			3	57.5	295.29	296.98	299.21	292.58	306.14	2.94
186	-19.19833333	146.8151667	Pacific	Central Indo-Pac Central and north Australia	Queensland			3	1	298.7	303.57	299.21	292.58	306.14	2.94
187	-19.19805556	146.815	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	1	298.7	303.57	299.21	292.58	306.14	2.94
189	-19.17066667	146.8465	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	0	295.84	296.47	299.09	292.57	306.11	2.92
190	-19.16944444	146.8472222	Pacific	Central Indo-Pac Central and north Australia	Queensland			2	0.5	301.47	300.76	299.09	292.57	306.11	2.92
193	-19.16916667	146.8469444	Pacific	Central Indo-Pac Central and north Australia	Queensland			5	4.75	297.02	300.69	299.09	292.57	306.11	2.92
194	-19.16916667	146.8469444	Pacific	Central Indo-Pac Central and north Australia	Queensland			4	3.5	294.83	296.25	299.09	292.57	306.11	2.92
198	-19.15430556	146.8660556	Pacific	Central Indo-Pac Central and north Australia	Queensland			1	20	295.18	296.98	299.07	292.2	305.77	2.88
199	-19.15430556	146.8660556	Pacific	Central Indo-Pac Central and north Australia	Queensland			3	6	298.66	304.27	299.07	292.2	305.77	2.88
201	-19.14833333	146.8702778	Pacific	Central Indo-Pac Central and north Australia	Queensland			3	4.5	297.14	300.69	299.07	292.2	305.77	2.88
202	-19.14833333	146.8702778	Pacific	Central Indo-Pac Central and north Australia	Queensland			6	8.25	297.14	300.69	299.07	292.2	305.77	2.88
203	-19.14833333	146.8702778	Pacific	Central Indo-Pac Central and north Australia	Queensland			4	2	294.89	295.33	299.07	292.2	305.77	2.88
204	-19.14833333	146.8702778	Pacific	Central Indo-Pac Central and north Australia	Queensland			4	2.25	294.89	295.33	299.07	292.2	305.77	2.88
206	-19.12269444	146.8813611	Pacific	Central Indo-Pac Central and north Australia	Queensland			3	6.25	298.66	304.27	299.07	292.2	305.77	2.88
207	-19.12188889	146.8808333	Pacific	Central Indo-Pac Central and north Australia	Queensland			3	5	295.18	296.98	299.07	292.2	305.77	2.88
209	-18.95605556	146.95625	Pacific	Central Indo-Pac Central and north Australia	Queensland			3	17.5	299.09	299.77	299.05	293.35	304.65	2.58
210	-18.91966667	-159.8466667	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			10	0.375	298.72	299.29	299.86	296.41	303.63	1.26
211	-18.88666667	-159.825	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			3	3.5	298.72	299.38	299.88	295.82	304.05	1.26
212	-18.88666667	-159.825	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			10	0.75	298.72	299.38	299.88	295.82	304.05	1.26
213	-18.85	-159.8083333	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			3	0.25	298.61	298.8	299.87	295.82	303.92	1.29
214	-18.85	-159.8083333	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			10	0	298.61	298.8	299.87	295.82	303.92	1.29
215	-18.85333333	-159.805	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			10	4	298.68	299.6	299.87	295.82	303.92	1.29
216	-18.85261111	-159.8061667	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			3	0.75	298.78	299.07	299.87	295.82	303.92	1.29
217	-18.85261111	-159.8061667	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			12	0.25	298.78	299.07	299.87	295.82	303.92	1.29
218	-18.84333333	-159.8025	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			3	17	298.78	299.07	299.87	295.82	303.92	1.29
219	-18.84333333	-159.8025	Pacific	Eastern Indo-Pac Cook Islands, so Cook Islands	Atitaki			10	0	298.78	299.07	299.87	295.82	303.92	1.29

Figura 2: Representación de la información guardada en la base de datos.

2.1. Muestreo

Tenemos una base datos con 10,000 entradas, estas serán tomadas como la población sobre la cual queremos sacar una muestra. Si tomamos los siguientes parámetros

- Margen: 10 %
- Nivel de confianza: 99 %
- Poblacion: 10000

y con la fórmula que revisamos en el curso; obtenemos que el tamaño ideal de la muestra debe ser 163.

Medidas	Valores
Promedio	6.41
Mediana	6
Moda	5
Desviacion std	3.43
Varianza	11.77

dataset_coral
Datos de interes
Tabla de frecuencia

Figura 3: Medidas de dispersión y tendencia central.

2.2. Medidas de dispersión y tendencia central

En la figura 3 están expresadas las medidas de tendencia central y dispersión, que revisamos en esta unidad, que fueron obtenidas de la muestra de la población. La muestra fue obtenida con ayuda de la funcion *muestra* de **Excel** y esta en otro libro porque son bastante datos.

De los datos podemos notar que el promedio y la mediana son bastante parecidas; esto implica que los datos estan parcialmente bien distribuidos, solo algo cargados en el percentil inferior. Aunque por otro lado, el que la moda no coincida con estos indica que hay valores *erráticos* en ambos extremos de los datos.

En otro análisis identificamos que la máxima profundidad registrada es 21,7, mientras que la mínima es 0,1.

Obtención de datos

En este trabajo se trabajo sobre el promedio, la mediana, la moda, la desviación estandar y la varianza. A continuación trataremos de definir estas variables y de

explicar como podemos obtenerlas.

Promedio Este se define como el *valor característico de una serie de datos cuantitativos*. Este se obtiene al dividir la sumatoria de todos los datos y dividirlos entre la cantidad de datos n que se tienen.

$$\text{promedio} = \bar{x} = \frac{\sum x_i}{n}$$

Para esto usamos la función **PROMEDIO** de excel, al cual sólo se le pasa el rango de los datos a promediar.

Mediana Por otro lado, esta representa el valor de la variable de posición central en un conjunto de datos ordenados; de manera que hay la misma cantidad de datos de ambos lados que esta entrada. En caso de que haya una cantidad par de datos, se saca un promedio de los dos datos centrales.

Esto también tiene una función en excel, **MEDIANA**. También sólo se le pasa el rango de los datos para sacar la mediana.

Moda Es el valor con mayor frecuencia en la distribución de datos. Esta es la única medida de tendencia central que se puede obtener para datos cualitativos y cuantitativos.

Para esta sigue extendiendo una fórmula ya definida en excel. Para la función **MODA** recibe los datos de los que se quiere obtener la moda.

Desviación estandar Es una medida de dispersión que se utiliza para cuantificar la variación o la dispersión de un conjunto de datos numéricos. La función en excel que la calcula es **STDEV** y recibe el rango de los datos a medir.

Varianza Es una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media. En excel se puede calcular de dos formas, como el cuadrado de la desviación estandar; o con la función **VAR**.

3. Conclusiones

Todas estas medidas y cálculos, que revisamos en la unidad, no permiten entender la manera en que los datos están esparcidos en el espacio. Sin estas herramientas deberíamos de tener la capacidad de leer los datos individualmente y tener la habilidad para contextualizar y recordar cada uno de los detalles de los datos. Por todo

esto, podemos concluir que estos cálculos definitivamente son útiles cuando se debe trabajar con grandes cantidades de datos cuantitativos.

Me parece que sería ideal que nos dieran más casos específicos y ejercicios en esta sección, esto nos permitiría entender mejor estos temas.

Referencias

- [1] Rivera C., B. (2020). *Evidencia de Aprendizaje U2*. No Publicado.
- [2] Universidad Abierta y a Distancia de México. (s/f). *Unidad 3. Representación numérica y gráfica de datos*. UnADM.
- [3] van Woesik, R. (2019). *Dataset: Global Bleaching and Environmental Data [Base de Datos]*. Florida Institute of Technology. <https://www.bco-dmo.org/dataset/773466>