



**Universidad Nacional Autónoma de México**  
**Escuela Nacional de Estudios Superiores**  
**Unidad Morelia**



Reporte Final  
**Análisis de Valores Nutricionales por Tipo de Dieta**

PRESENTA:  
**Alexis Uriel Aguilar Uribe**

PROFESORES:  
**Dra. María Del Río Francos**  
**Dr. César Andrés Torres Miranda**

GRADO  
**Licenciatura en Tecnologías para la Información en Ciencias**

**Asignatura:** Estadística Descriptiva e Inferencial

**A:** 26 de Mayo del 2025

# Índice

<b>Anexos</b>	<b>1</b>
<b>1. Introducción</b>	<b>2</b>
<b>2. Objetivos Generales</b>	<b>2</b>
<b>3. Presentación de los Datos</b>	<b>3</b>
3.1. Fuente de Datos . . . . .	3
3.2. Interés del Estudio . . . . .	3
3.3. Variables del Conjunto de Datos . . . . .	3
3.4. Ejemplo de Registros en el Conjunto de Datos . . . . .	4
<b>4. Estadística Descriptiva</b>	<b>5</b>
4.1. Preprocesamiento (Transformación) de los Datos . . . . .	5
4.2. Descripción de los Valores de las Variables . . . . .	6
4.3. Visión General de los Datos . . . . .	6
4.3.1. Dieta DASH . . . . .	9
<b>5. Análisis Bivariado</b>	<b>11</b>
<b>6. Muestreo e Intervalos de Confianza</b>	<b>12</b>
<b>7. Pruebas de Hipótesis</b>	<b>13</b>
<b>A. Marco Teórico</b>	<b>13</b>
A.1. DASH (Dietary Approaches to Stop Hypertension) . . . . .	13
A.2. Dieta Keto . . . . .	13
A.3. Dieta Mediterránea . . . . .	14
A.4. Dieta Paleo (Paleolítica) . . . . .	14
A.5. Dieta Vegana . . . . .	14
<b>Referencias Bibliográficas</b>	<b>15</b>

# 1. Introducción

Este trabajo tiene como fin de exponer el proceso llevado a cabo para realizar el análisis estadístico de los valores nutricionales (macronutrientes) que aportan las dietas: *DASH* (Dietary Approaches to Stop Hypertension), *keto*, *mediterránea*, *paleo* (paleolítica) y *vegana*.

Siendo el principal enfoque el responder si hay una diferencia nutricional significativa entre las diferentes dietas. En decir, hacer uso de técnicas de estadística descriptiva e inferencial para probar si existe una diferencia en los aportes nutricionales entre las distintas dietas que están siendo estudiadas. La anterior prueba se basa en recetas de diferentes cocinas a nivel mundial, y sobre éstas últimas serán auxiliares para realizar un estudio más granulado sobre el comportamiento de las dietas en escenarios más específicos.

# 2. Objetivos Generales

Para la realización de lo anterior expuesto, se puntualizan los objetivos del proyecto:

- Realizar de un análisis estadístico de los macronutrientes en las diferentes dietas con el fin de caracterizar sus aportes nutricionales y sus distinciones en las diferentes cocinas.
- Conjeturar y probar hipótesis relacionadas a preguntas de interés sobre los aportes nutricionales en cada dieta en base al análisis estadístico.
- Probar si existe una diferencia significativa en los aportes nutricionales entre las diferentes dietas con el fin de probar si cada dieta de estudio es única.

## **3. Presentación de los Datos**

### **3.1. Fuente de Datos**

El conjunto de datos con el que se está trabajando para este proyecto se encuentran en [1], publicado por la comunidad de Kaggle. Los datos consisten de un conjunto de recetas de diferentes dietas y cocinas, además incluye información de los macronutrientes que aporta cada receta.

[1] Aunque en la descripción ni en los metadatos del conjunto de datos se haga mención de las fuentes explícitas de los datos ni el objetivo de esta extracción, sí cuenta con una sección de cómo usar el conjunto de datos, ideas de investigación y reconocimientos.

De los apartados de cómo usar el conjunto de datos e ideas de investigación, se encuentra una idea, implícita, de la información que se quería estudiar. La principal información de interés se vuelve que es: el crear planes alimenticios saludables, ya sea usando las recetas proporcionadas o creando unas nuevas basadas en una dieta y cocina, y el estudiar la relación entre dieta y salud.

Del apartado de reconocimientos, se concluye que las recetas fueron proporcionadas por diferentes creadores de las mismas y demás contribuidores al conjunto de datos.

### **3.2. Interés del Estudio**

Se consultó [2] en sus capítulos 4 y 8, de donde se proporciona un mejor entendimiento de la importancia de los macronutrientes y una descripción general de las dietas en este trabajo, resultando interesante que en cada dieta se consumen diferentes alimentos y productos con ciertas características para ya sea respetar alguna creencia, fundamento o cuota de macronutrientes. De esto último, proporciona un indicio de que existe una diferencia entre las dietas a nivel de sus aportes nutricionales, por lo tanto, lo que se quiere realizar es probar esta diferencia de manera significativa haciendo uso de la estadística y, en caso de que la haya, mostrar que tanta es ésta diferencia y sus implicaciones.

### **3.3. Variables del Conjunto de Datos**

El conjunto de datos consta de las siguientes variables. Se menciona su nombre, el tipo de variable y sus valores (en total y únicos):

Variable	Nombre	Tipo	Cantidad de Datos	Valores Únicos
1	Diet_type	Cualitativa Nominal	7806	5
2	Recipe_name	Cualitativa Nominal	7806	7062
3	Cuisine_type	Cualitativa Nominal	7806	19
4	Protein	Cuantitativa Continua	7806	6060
5	Carbs	Cuantitativa Continua	7806	6618
6	Fat	Cuantitativa Continua	7806	6322

La variable *Recipe\_Name* no es relevante para este trabajo pero figura dentro del dataset. Se hace mención que el conjunto de datos no presenta valores faltantes.

### 3.4. Ejemplo de Registros en el Conjunto de Datos

Para ejemplificar como luce el conjunto de datos, se presente una instancia de cada tipo de dieta:

Diet_type	Recipe_name	Cuisine_type
dash	Old Fashioned	world
keto	Keto Egg Drop Soup	chinese
mediterranean	Mediterranean Mix	mediterranean
paleo	Easy Paleo Herb Gravy recipes	french
vegan	Braised Green Beans with Tomatoes	mediterranean

Protein	Carbs	Fat
0.12	9.66	0.02
21.31	9.11	60.88
8.11	9.59	14.64
23.56	39.05	42.25
17.49	77.86	70.20

## 4. Estadística Descriptiva

### 4.1. Preprocesamiento (Transformación) de los Datos

De las instancias presentadas en Subsección 3.4 Ejemplo de Registros en el Conjunto de Datos, se tiene que los valores de los macronutrientes pueden tomar un amplio rango de valores, esto puede generar un conflicto al momento de generar una comparativa entre dietas, debido a principalmente los rangos de valores. Para resolver esta situación los valores de los macronutrientes son normalizados con la norma  $l1$ , es decir, se calcula el total de macronutrientes (que se guarda como otra variable en *Total\_macronutrients*) de cada receta y cada macronutriente se divide por este total.

	dash	keto	mediterranean	paleo	vegan	Suma Dietas
kosher	5	0	0	2	0	7
caribbean	3	7	1	6	1	18
central europe	9	11	1	9	4	34
japanese	9	10	2	5	24	50
eastern europe	10	11	3	27	4	55
middle eastern	21	17	26	12	15	91
indian	20	12	3	9	48	92
chinese	38	38	1	26	17	120
asian	24	11	12	12	67	126
south american	54	21	10	21	31	137
south east asian	31	34	8	29	46	148
nordic	32	35	31	45	9	152
mexican	61	60	17	48	38	224
british	64	90	4	54	27	239
world	234	6	6	3	10	259
french	150	163	61	154	76	604
italian	165	234	148	171	81	799
mediterranean	176	89	1274	106	99	1744
american	639	663	145	535	925	2097
Suma Cocinas	1745	1512	1753	1274	1522	7806

Al considerar la cantidad de recetas que hay por dieta y cocina, se descubre que, para ciertas grupos o configuraciones no contienen recetas, por lo que para mitigar esta falta de instancias lo que se realizar es juntar los *Cuisine\_type* en base a la cercanía geográfica, esto debido a que son colindantes comparten historia, cultura y, lo más relevante, ideas gastronómicas. Por ello, las *Cuisine\_type* se reagrupan de la siguiente manera:

Grupo de <i>Cuisine_type</i>	<i>Cuisine_type</i>
american	american
mediterranean	mediterranean
world	world
latin american	mexican, south american, caribbean
european	italian, french, nordic, eastern europe, central europe, kosher, british
asian	chinese, indian, south east asian, middle eastern, asian, japanese

## 4.2. Descripción de los Valores de las Variables

Para el presente trabajo se harán uso de las siguientes variables, se acompañan con una descripción de su significado o representación:

- **Diet\_type:** Variable nominal que representa el tipo de dieta (DASH, keto, mediterránea, paleo, vegana) a la que pertenece una receta. Con esta variable se va permitir estratificar las recetas y estudiarlas de una manera más granular, y para realizar hipótesis sobre lo qué está pasando en una dieta o entre las diferentes dietas.
- **Cuisine.type:** Variable nominal que representa a qué (estilo de) cocina o región (mexicana, americana, italiana, entre otras) pertenece una receta. Al usarla va a permitir el comparar cómo son las recetas de una dieta en diferentes regiones, y realizar comparativas a lo largo de las dietas.
- **Protein:** Variable continua que representa el porcentaje, respecto al total de macronutrientes, de proteínas que son aportados por una receta.
- **Carbs:** Variable continua que representa el porcentaje, respecto al total de macronutrientes, de carbohidratos que son aportados por una receta.
- **Fat:** Variable continua que representa el porcentaje, respecto al total de macronutrientes, de grasas que son aportados por una receta.
- **Total\_Macronutrients:** Variable continua que representa el total de macronutrientes que son aportados por una receta. Esta variable es auxiliar para la prueba de hipótesis, y no estará presente en la parte del Análisis Estadístico,

## 4.3. Visión General de los Datos

Primero se presenta un análisis sobre los macronutrientes de las recetas sin estratificarlas según el tipo de dieta:

Medida	Carbs	Protein	Fat
Media	0.433471	0.234762	0.331767
$Q_1$	0.205251	0.110188	0.184583
$Q_2$	0.432028	0.190931	0.314359
$Q_3$	0.635058	0.338059	0.464532
Desviación Estándar	0.256032	0.163886	0.194920
Mínimo	0.000330	0.000000	0.000000
Máximo	1.000000	0.887557	0.997940
Asimetría de Fisher	0.189556	0.922401	0.461455

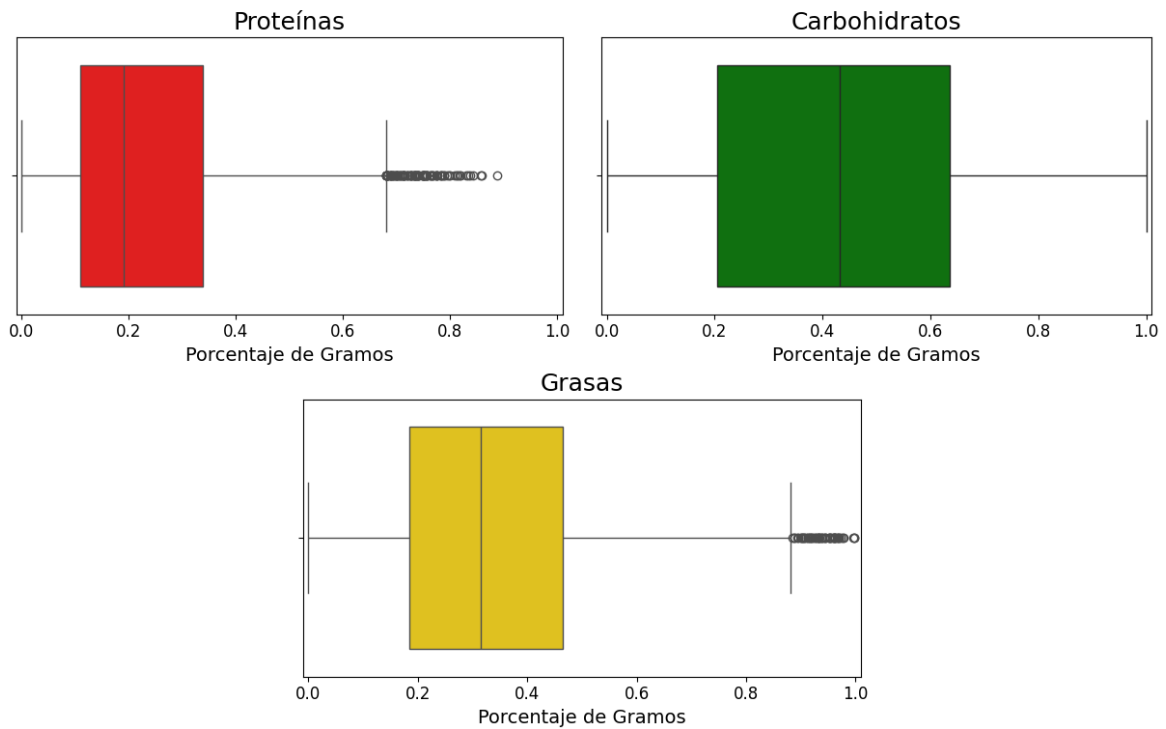
Se reportan bajos valores en proteínas en comparación con los carbohidratos y grasas si se hace uso de la mediana ( $Q_2$ ), dicho así: el cincuenta por ciento de las recetas tienen a lo mucho 19,09 % de proteínas, en comparación con el 43,20 % de carbohidratos y el 31,43 % de grasas. Esto es un indicio de que las recetas, en general, tienden a ser altas en carbohidratos y grasas entre las diferentes dietas y cocinas; mientras que son bajas en proteínas. Este último punto se apoya al considerar la asimetría que reportan las proteínas a ser un valor extremadamente alto y positivo.

Si se gráfica la distribución de los macronutrientes se tiene que, debido a la asimetría y a la desviación estándar, contienen datos atípicos en proteínas y grasas en una región positiva respecto a la mediana, y esto se relaciona con lo mencionado de que una receta no tiende a un aporte alto de proteínas ni de grasas. Y si se consider el rango intercuartil, se observa que en estos macronutrientes es menor, en comparación, que con el de los carbohidratos. Por ello, se tiene que una receta es rica en carbohidratos y este patrón se repite en todas las dietas por las asimetrías y desviación estándar.

	dash	keto	mediterranean	paleo	vegan
american	639	663	145	535	925
asian	143	122	52	93	217
european	435	544	248	462	201
latin american	118	88	28	75	70
mediterranean	176	89	1274	106	99
world	234	6	6	3	10

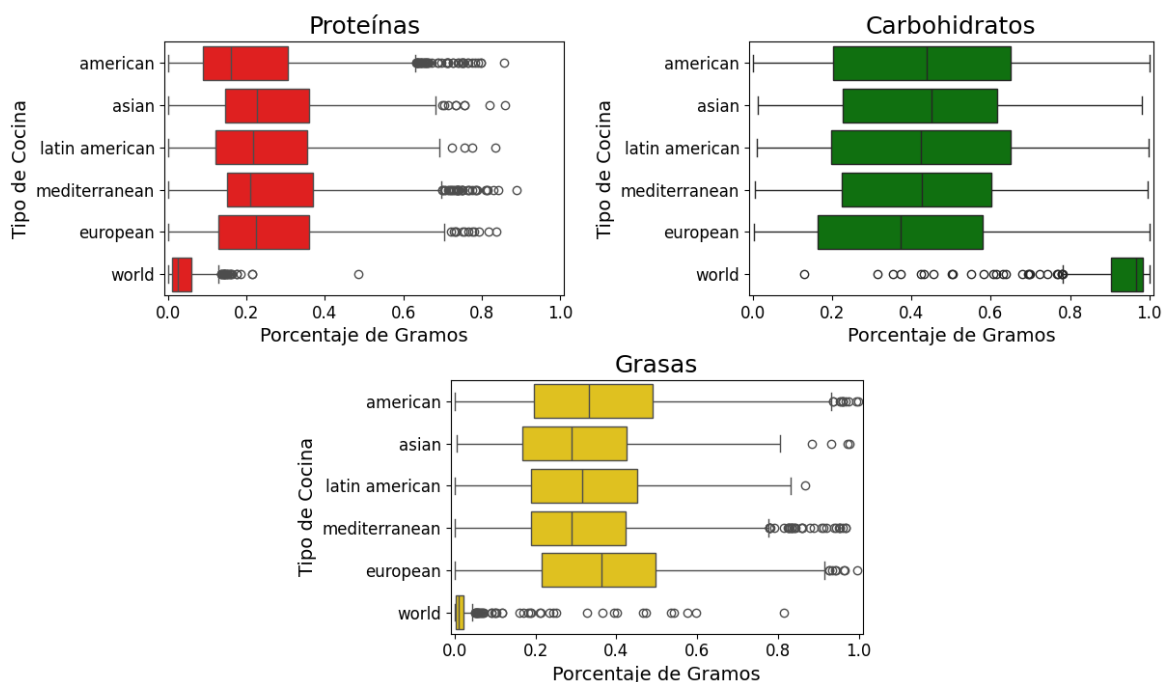


## Distribución de Macronutrientes



Al considerar la influencia de *Cuisine\_type*, se tiene que las diferentes cocinas conservan ligeros matices sobre la distribución de los macronutrientes, es decir, pero por el solapamiento de las cajas y bigotes se podría esperar que no tengan una diferencia significativa, indicando que las cocinas no tienen una influencia notoria sobre los macronutrientes. En el tipo de cocina *world* parece ser que tiene un comportamiento atípico en comparación con las demás cocinas, por ello se va a descartar del análisis.

## Distribución de Macronutrientes Por Cocina

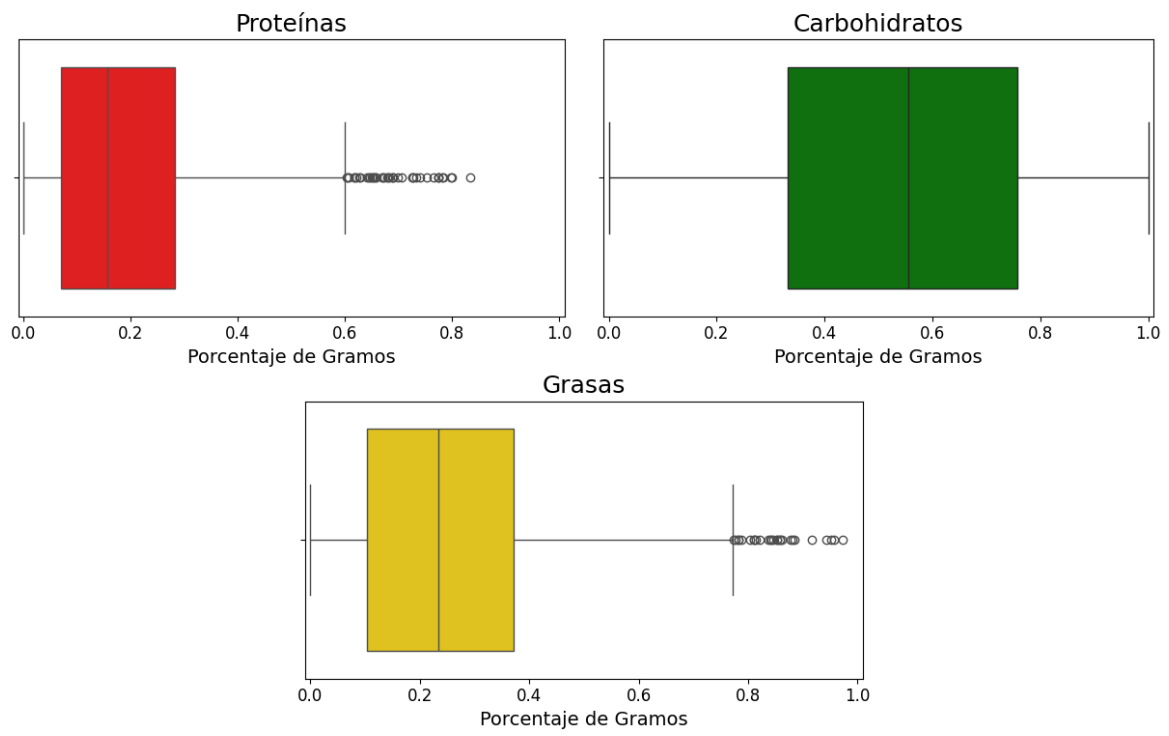


#### 4.3.1. Dieta DASH

Una receta de esta dieta tendrá que, en promedio, el 55 % de sus macronutrientes son carbohidratos (provenientes de frutas, vegetales y granos enteros); el 25 % son grasas que, por su naturaleza, son saludables; y el 20 % son proteínas, las cuáles provienen de carnes magras. Aunque esta dieta se menciona ser saludable para la salud cardiovascular, no implica que exista un balance o equilibrio en los macronutrientes consumidos por receta. Debido a la desviación estándar y rango intercuartil de las de proteínas y grasas, se tiene que estos macronutrientes se encuentran concentrados en un rango más pequeño de valores en comparación con los carbohidratos. De lo mencionado, podría significar que la contribución de los macronutrientes no son tan variadas como lo que se esperaría contradiciendo que sea una dieta saludable, notando que es una dieta rica en carbohidratos. Esto no excluye que el consumir varias recetas (comidas) se logre un balance.

Medida	Carbs	Protein	Fat
Media	0.549425	0.196241	0.254334
$Q_1$	0.331143	0.068931	0.103381
$Q_2$	0.555219	0.156626	0.234742
$Q_3$	0.757917	0.282629	0.371292
Desviación Estándar	0.278850	0.162871	0.194078
Mínimo	0.001526	0.000000	0.000000
Máximo	1.000000	0.833467	0.973404
Asimetría de Fisher	-0.057984	1.101171	0.732534

# Distribución de Macronutrientes



## 5. Análisis Bivariado

## 6. Muestreo e Intervalos de Confianza

## 7. Pruebas de Hipótesis

### A. Marco Teórico

La dieta es uno de los principales factores de riesgo de las enfermedades crónicas, y las enfermedades sensibles a la dieta contribuyen en gran medida a los costes sanitarios mundiales. Se han propuesto literalmente miles de *dietas*, que pueden describirse en términos generales como basadas en creencias, en alimentos específicos o en nutrientes; centradas en la pérdida de peso o en el aumento de peso (muscular); dietas de desintoxicación (detox) y dietas diseñadas por razones médicas específicas.[2]

Las *dietas de moda* son dietas populares durante un tiempo sin basarse necesariamente en una recomendación dietética estándar. A menudo promueven una pérdida de peso irracionalmente rápida o afirmaciones de salud sin sentido, y se anuncian como dietas que requieren poco esfuerzo por parte de quien las sigue. La promesa de ganancias fáciles, combinada con la presión social para lograr un determinado tipo de cuerpo, puede dejar al público susceptible a afirmaciones infundadas o exageradas.[2]

Las dietas estudiadas desde una perspectiva estadística en el presente trabajo, son englobadas en las *dietas de moda*, que a veces son referidas como *dietas sin evidencia científica*. Siendo la dieta DASH la única que cuenta con algún tipo de fundamento.

#### A.1. DASH (Dietary Approaches to Stop Hypertension)

[2] La dieta DASH (Enfoques Dietéticos para Detener la Hipertensión) es un patrón dietético diseñado específicamente para ayudar a reducir la presión arterial y promover la salud general del corazón. Hace hincapié en el consumo de una variedad de alimentos ricos en nutrientes, como frutas, verduras, cereales integrales, proteínas magras y productos lácteos bajos en grasa, y en la limitación de la ingesta de sodio, grasas saturadas y azúcares añadidos.

#### A.2. Dieta Keto

[2] Una dieta baja en hidratos de carbono (baja en carbohidratos) es un patrón alimentario que restringe la ingesta de carbohidratos, sustituyéndolos normalmente por mayores cantidades de proteínas y grasas. La dieta cetogénica es una forma de dieta baja en carbohidratos con un alto contenido en grasas en relación con la ingesta de proteínas y carbohidratos.

El objetivo de la dieta cetogénica es inducir la cetosis, un estado metabólico que se produce cuando el cuerpo quema grasa para obtener energía en lugar de glucosa, lo que induce la pérdida de peso.

### **A.3. Dieta Mediterránea**

[2] La dieta mediterránea es un patrón alimentario inspirado en los hábitos alimenticios tradicionales de los países situados a orillas del mar Mediterráneo. Se caracteriza por un alto consumo de frutas, verduras, cereales integrales, legumbres, frutos secos y aceite de oliva; un consumo moderado de pescado y aves; y un bajo consumo de carnes rojas, alimentos procesados y dulces.

### **A.4. Dieta Paleo (Paleolítica)**

[2] La dieta paleo, también conocida como dieta paleolítica o dieta del hombre de las cavernas, es un enfoque dietético que pretende imitar los hábitos alimentarios de nuestros antiguos antepasados del Paleolítico.

Hace hincapié en el consumo de alimentos integrales y no procesados que habrían estado al alcance de los primeros humanos, como carnes magras, pescado, frutas, verduras, frutos secos y semillas, y excluye los cereales, las legumbres, los productos lácteos, los alimentos procesados y los azúcares añadidos.

### **A.5. Dieta Vegana**

[2] La dieta vegana es un patrón dietético basado en plantas que excluye el consumo de todos los productos de origen animal. Se centra en el consumo de una variedad de alimentos de origen vegetal, como frutas, verduras, cereales legumbres, frutos secos y semillas.

Es importante señalar que, aunque las dietas veganas pueden ser nutricionalmente adecuadas, debe prestarse atención a garantizar una ingesta suficiente de nutrientes esenciales como proteínas, hierro, calcio, vitamina B12 y ácidos grasos omega-3.

## Referencias Bibliográficas

- [1] T. D. [www.kaggle.com/thedevastator], *Diets, Recipes And Their Nutrients*, kaggle, 2024.  
dirección: <https://www.kaggle.com/datasets/thedevastator/healthy-diet-recipes-a-comprehensive-dataset>.
- [2] F. F. Marvasti, «Popular Diets and Health,» *Culinary Medicine*,