

# Desafío: Diagrama Entidad-Relación

## Tabla de contenido

- [Desafío: Diagrama Entidad-Relación](#)
  - [Tabla de contenido](#)
  - [Consideraciones](#)
  - [Dataset](#)
  - [Diagrama Entidad-Relación](#)
    - [Relaciones y atributos](#)

## DIAGRAMA ENTIDAD RELACIÓN

**Formato:** PPT o gSlides

**Sugerencia:** En caso de ser un archivo en línea, activar permisos de acceso.

Desafío  
entregable



### >> Consigna:

- Diseñar el modelo entidad-relación de al menos dos de las temáticas elegidas para el proyecto final.

### >>Aspectos a incluir en el entregable:

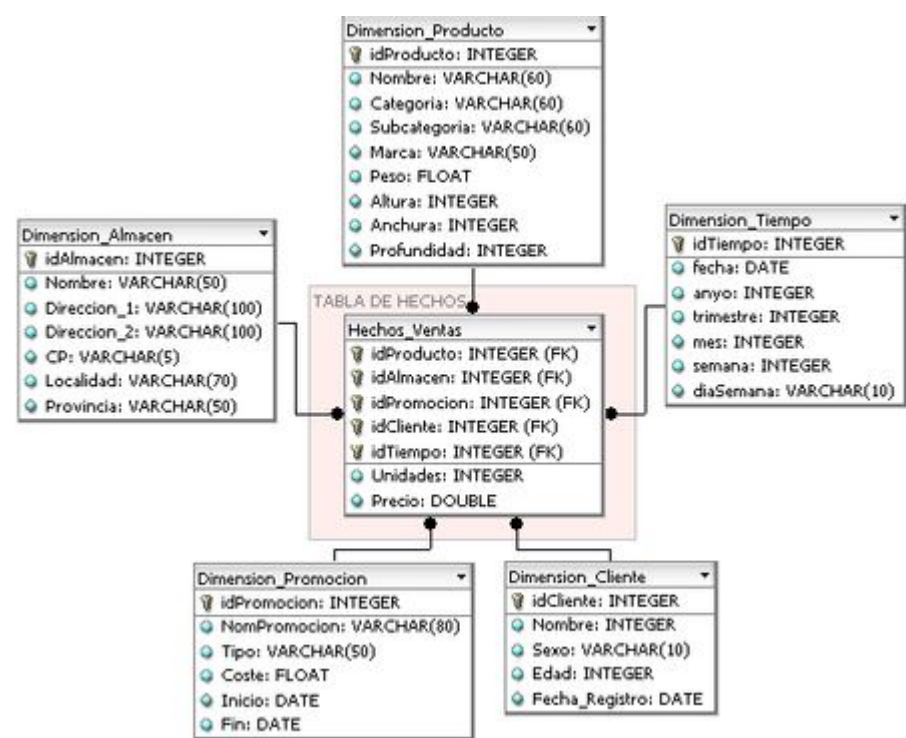
- Definir al menos cinco tablas.
- Crear el diagrama de entidad-relación con todos sus componentes:
  - Entidades.
  - Acciones de relacionamiento.
  - Tipos de relación.
  - Campos clave.

## Consideraciones

En el presente desafío se realizó inicialmente el análisis de los datos que comprendían el dataset elegido, en este caso, "Data Science Jobs Salaries", el cual fue encontrado en el repositorio *Kaglee* al cual se puede acceder a través del siguiente [link](#)

**Esquema en estrella:** Según wikipedia, \*En las bases de datos usadas para data warehousing, **un esquema en estrella** es un modelo de datos que tiene una **tabla de hechos** (o tabla fact) que contiene los datos para el análisis, rodeada de las **tablas de dimensiones**. Este aspecto, de tabla de hechos (o central) más grande rodeada de radios o tablas más pequeñas es lo que asemeja a una estrella, dándole nombre a este tipo de construcciones. Las tablas de dimensiones tendrán siempre una **clave primaria** simple, mientras que en la tabla de hechos, la clave principal estará compuesta por las claves principales de las tablas dimensionales.

Ejemplo de modelo de datos en estrella:



Fuente: [https://es.wikipedia.org/wiki/Esquema\\_en\\_estrella](https://es.wikipedia.org/wiki/Esquema_en_estrella)

## Dataset

El dataset inicial se encontraba en formato `.csv` separado por comas, y fue seteado de forma que tal que fue separado en 5 tablas principales, las cuales fueron usadas para el diseño del diagrama Entidad-Relación.

	A	B	C	D	E	F
1	work_year,	experience_level,	employment_type,	job_title,	salary,	salary_currency,
2	2021e,EN,FT,	Data Science Consultant,	54000,EUR,	64369,DE,	50,DE,	L
3	2020,SE,FT,	Data Scientist,	60000,EUR,	68428,GR,	100,US,	L
4	2021e,EX,FT,	Head of Data Science,	85000,USD,	85000,RU,	0,RU,	M
5	2021e,EX,FT,	Head of Data,	230000,USD,	230000,RU,	50,RU,	L

El `dataset final` fue diseñado a partir de 5 tablas extraídas y editadas del dataset original, entre otras, jobs, remote\_ratio, company, salary y employee. Por último, el modelo de datos se diseñó con un esquema en estrella.

## Diagrama Entidad-Relación

Para el diseño del presente diagrama se tuvo en cuenta que la tabla jobs se consideró como la tabla hechos por tratarse del tema principal en cuanto a los salarios de empleos en áreas de Data Science, las demás fueron clasificadas como tablas de dimensiones. El diagrama se realizó en el programa de `Microsoft Visio` y se exportó en formato `.jpg`:



Relaciones y atributos

- Tabla "job":

	A	B	C	D	E	F	G
1	PK_id_job	job_title	work_year	FK_id_ratio	FK_id_company	FK_id_salary	FK_id_employee
2		1 Data Science Consultant	2021e	2	1	1	1
3		2 Data Scientist	2020	3	2	2	2
4		3 Head of Data Science	2021e	1	3	3	3
5		4 Head of Data	2021e	2	4	4	4

- Tabla "remote\_ratio":

	A	B	C
1	PK_id_ratio	remote_ratio	ratio_label
2		2	50 Partially remote
3		3	100 Fully remote
4		1	0 No remote work
5		2	50 Partially remote

- Tabla "company"

	A	B	C
1	PK_id_company	company_location	company_size
2	1	DE	L
3	2	US	L
4	3	RU	M
5	4	RU	L
6	5	US	S

- Tabla "salary"

	A	B	C	D
1	PK_id_salary	salary	salary_currency	salary_in_usd
2	1	54000	EUR	64369
3	2	60000	EUR	68428
4	3	85000	USD	85000
5	4	230000	USD	230000
6	5	125000	USD	125000

- Tabla "employee"

	A	B	C	D	E	F
1	PK_id_employee	experience_level	experience_label	employment_type	employment_label	employee_residence
2	1	EN	Entry-level / Junior	FT	Full-time	DE
3	3	SE	Senior-level / Expert	FT	Full-time	GR
4	4	EX	Executive-level / Director	FT	Full-time	RU
5	4	EX	Executive-level / Director	FT	Full-time	RU