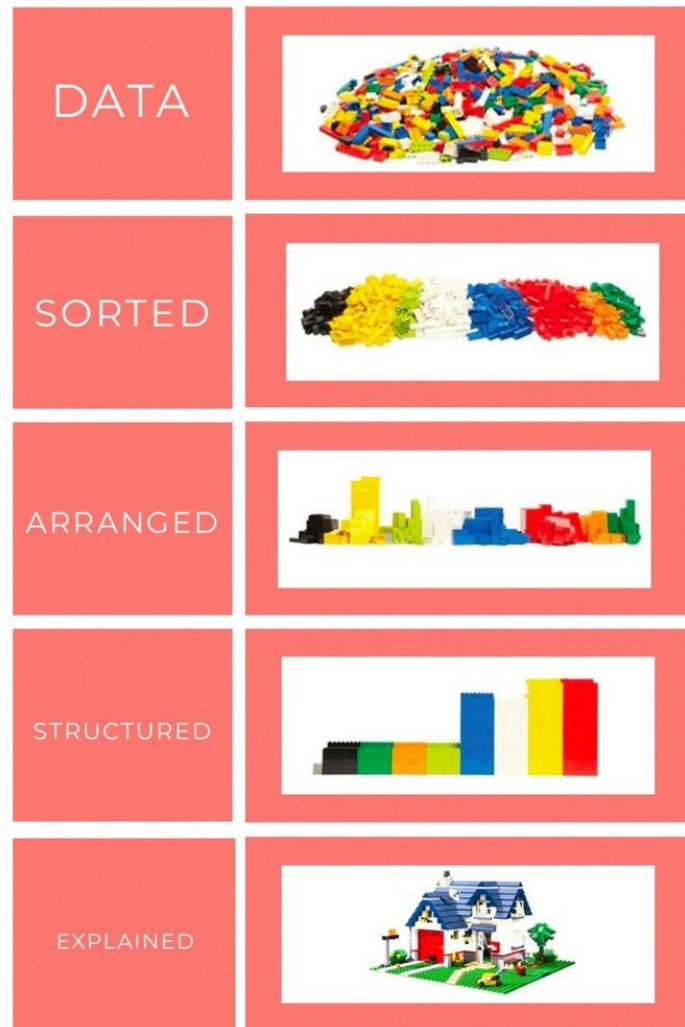


# CIENCIA DE DATOS

GRADO EN BIOMEDICINA

DR. ARIEL CARIAGA-MARTÍNEZ

# Fases y herramientas en Ciencia de Datos



¿De qué forma recogemos los datos?

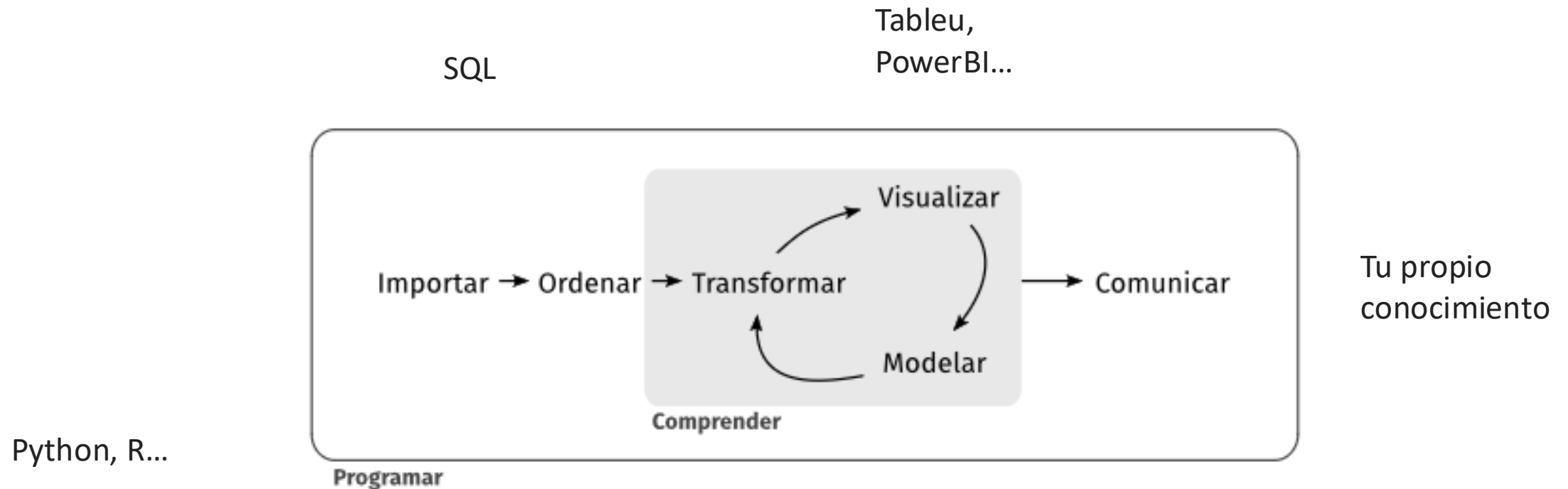
¿Cómo los podríamos organizar?

¿Cómo los podríamos transformar?

¿Cómo los podríamos modelar?

¿QUÉ HISTORIA (CONVINCENTE) PODEMOS CONTAR?

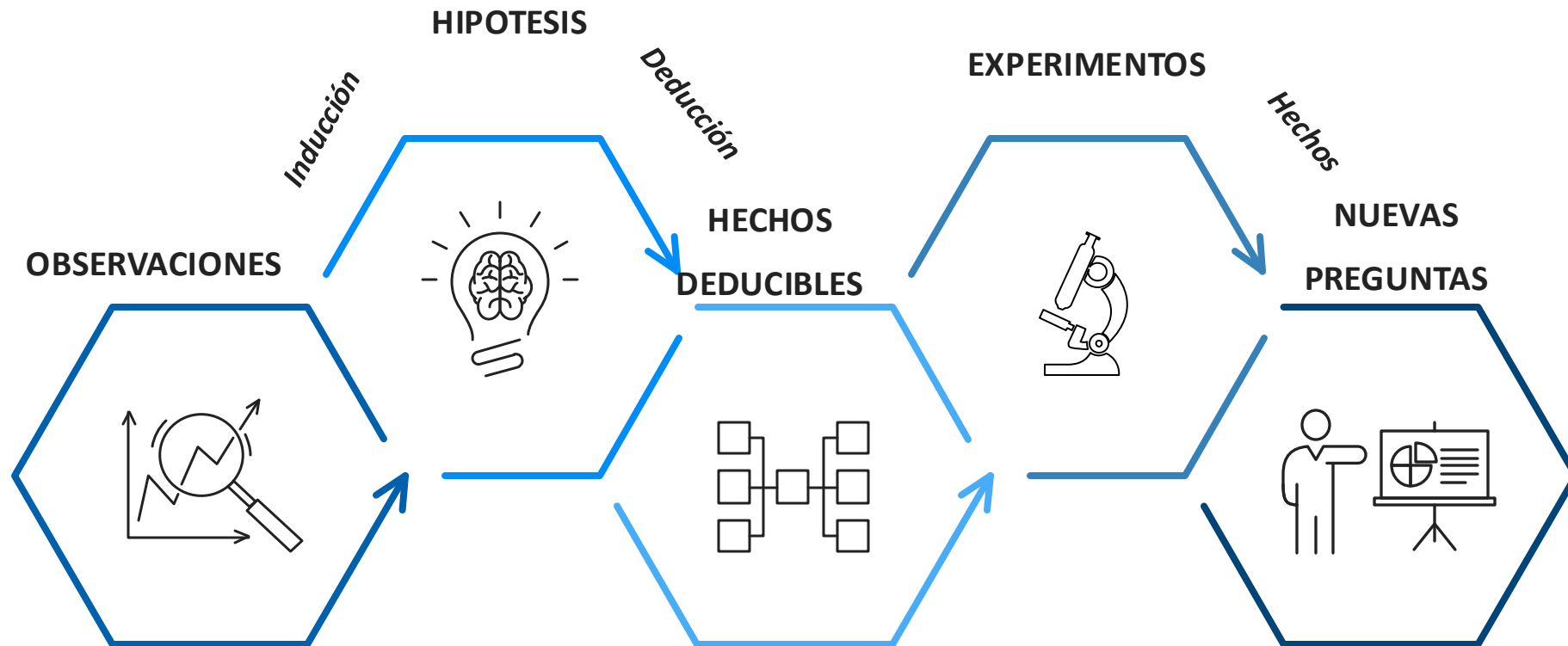
# Fases y herramientas informáticas en Ciencia de Datos.



ETL/ELT: extract, (load) transform, load

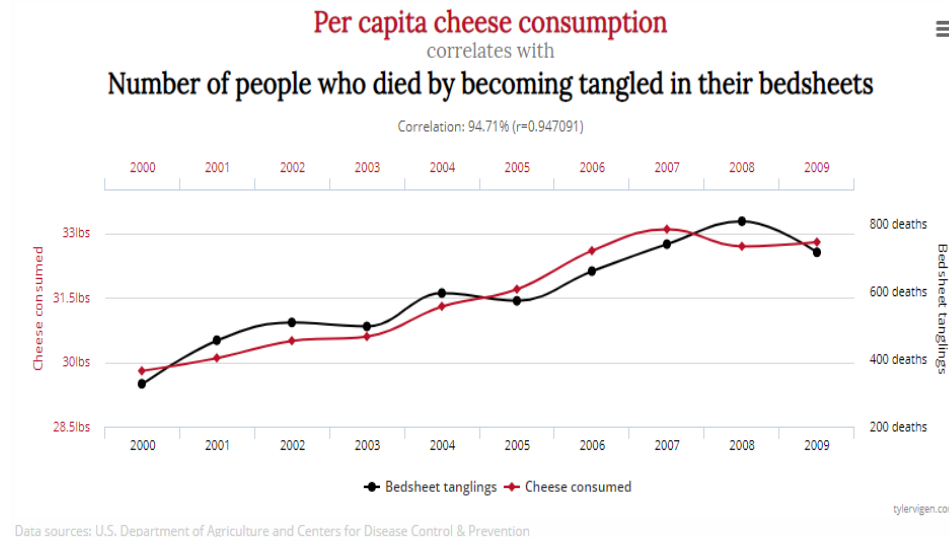
<https://es.r4ds.hadley.nz>

## INICIO=PROBLEMA



# Preguntas clave

- ¿Qué pregunta quiero responder?
- ¿A qué tipos de datos tengo acceso? (estructurados, semi-estructurados, no estructurados.
- ¿A qué cantidad de datos tengo acceso?
- **¿QUÉ CONOCIMIENTO TENGO DEL NEGOCIO?**
- ¿A qué herramientas tengo acceso? → No debería ser un problema (técnica vs tecnologías).



Data sources: U.S. Department of Agriculture and Centers for Disease Control & Prevention

# GRACIAS

**DR. ARIEL CARIAGA-MARTINEZ**

CIENCIA DE DATOS

ACARIMAR@UAX.ES

# TRABAJOS

GRADO EN BIOMEDICINA

DR. ARIEL CARIAGA-MARTÍNEZ

# UN “HECHO CIENTÍFICO”

## Un “hecho científico”

Siguiendo con los **resultados/datos** del trabajo anterior:

- Formar equipos (2-4 personas).
- Seguir el gráfico del método científico y (RE)aplicar cada fase CONSIDERANDO los resultados previamente obtenidos.
- **AHORA EL DATASET ESTÁ ACTUALIZADO.**
- Por ejemplo: ¿qué observaciones surgieron? A partir de la/s conclusión/es ¿se pueden proponer otras hipótesis? ¿Podríamos obtener “hechos deducibles” desde las hipótesis planteadas? ¿Qué “experimentos” podríamos proponer?
- EL OBJETIVO SERÁ OBTENER UN INFORME RAZONADO DE LOS PASOS SEGUIDOS EN UN MÁXIMO DE 2 FOLIOS (NO NECESARIAMENTE “ESCRITO”: IMÁGENES, GRÁFICOS, FLUJOS). DEBE SER **AUTOEXPLICATIVO** Y SE PUEDE REALIZAR UN “PRIMER ANÁLISIS GENERAL”.

Pueden ser opciones  
“teóricas” o “prácticas”.

## PENSAMIENTO CRÍTICO