



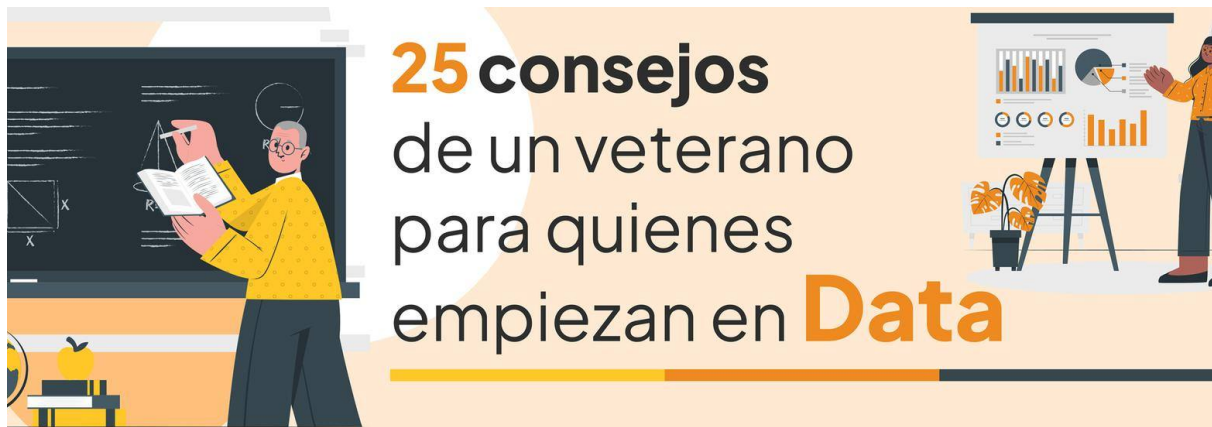
Universidad de Buenos Aires
Facultad de Ciencias Económicas



Consejos para empezar con los datos



Material de Estudio



25 consejos de un veterano para quienes empiezan en Data



01. Aprende SQL como si fuera tu idioma nativo
Es el inglés del dato: todos lo entienden, todos lo usan.



10. No te olvides de la seguridad
Gobernanza, accesos, PII, GDPR... mejor ser paranoico desde el inicio.



19. Usa entornos de prueba
Nunca experimentes con datos reales de producción.



02. No te cases con una herramienta
Hoy manda Databricks, mañana Snowflake, pasado Fabric. Lo que no cambia es la lógica del dato.



11. Automatiza los procesos repetitivos
Si haces lo mismo tres veces, métele en un script.



20. Pregunta al equipo senior
Más vale parecer pesado que repetir errores ya resueltos.



03. Versiona todo:
Código, queries, notebooks, incluso tus pipelines. Git será tu red de seguridad.



12. Cuida tu naming convention
"tabla1_final_v3_definitiva_OK" es la semilla del caos.



21. Aprende sobre arquitecturas modernas:
Lakehouse, Data Mesh, Streaming. No te quedes solo en el batch clásico.



04. Data ≠ Excel
Aprende a pensar en pipelines, no en hojas sueltas.



13. Aprende un lenguaje de scripting (Python o Scala)
SQL es vital, pero no suficiente.



22. Cuidado con el hype
No todo necesita IA generativa ni un modelo de deep learning. A veces un GROUP BY basta.



05. Entiende el negocio antes de picar código
Un dashboard bonito que no responde preguntas reales no sirve de nada.



14. No subestimes el poder de un buen dashboard
Pero ojo: menos es más. KPIs claros y directos.



23. Practica el "menos es más"
Pipelines simples son más fáciles de mantener que obras de arte imposibles de entender.



06. Los datos sucios son la norma, no la excepción
Asumelo: el 70% de tu tiempo se irá en limpiar y validar.



15. DataOps y MLOps no son buzzwords
Son lo que diferencia un experimento de un sistema productivo estable.



24. No olvides las soft skills
Explicar un modelo de chum a un director de marketing es tan importante como entrenarlo bien.



07. Haz logs y métricas de todo
Sin observabilidad, tus pipelines son cajas negras que nadie puede arreglar.



16. Conoce bien los costes de la nube
Escalar está bien, pero apagar el cluster a tiempo es aún mejor.



25. Ten paciencia y humor
Los datos siempre fallan en viernes a las 6 de la tarde. Respira: es parte del oficio.



08. Pregunta siempre "¿y este dato de dónde viene?"
El lineage importa más de lo que crees.



17. Juega con datos públicos
Kaggle, Open Data, APIs... ahí se entrena el músculo sin romper nada en producción.



09. Aprende a optimizar queries
Un JOIN mal hecho puede costar miles en la nube.



18. Aprende a documentar
Nadie lee manuales eternos, pero un buen README salva vidas.



Material de Clases

Compilado por **Aníbal M. Mazza Fraquelli** Doctor de la Universidad de Buenos Aires para el uso de sus clases en la Facultad de Ciencias Económicas de la Universidad de Buenos Aires.

Contenidos de esta página

Los contenidos **aquí incluidos integran desarrollos y escritos propios del autor, así como materiales de terceros (documentos, textos, fragmentos, conceptos, imágenes, esquemas, definiciones u otros recursos)**, los cuales son utilizados a título ilustrativo, explicativo o formativo, respetando la normativa vigente en materia de derechos de autor y citando las fuentes cuando corresponde.

La selección, organización, adaptación pedagógica y contextualización de los contenidos constituye un trabajo original del autor, orientado a facilitar los procesos de enseñanza y aprendizaje.

Este material no persigue fines comerciales y su reproducción, total o parcial, queda limitada al ámbito educativo, debiendo preservarse siempre la mención de la autoría y las fuentes originales.

Autorización de uso

Se permite la reproducción, comunicación pública, distribución y utilización total o parcial de los contenidos de su material, en formato físico o digital, con fines exclusivamente educativos, académicos o de divulgación, siempre que se respete la integridad del contenido y se incluya la correspondiente referencia a la fuente y a la autoría.

Las ideas, opiniones e interpretaciones contenidas en este material corresponden exclusivamente al autor.

Queda expresamente excluido cualquier uso con fines comerciales.