PL-300 · CHEAT SHEET · CERTIFYDECK

Power BI Data Analyst

Dominios, pesos oficiales y habilidades clave del temario de PL-300, para repasar antes del examen.

Nivel: Intermediate (Associate) Duración: 100 min Nota de corte: 700/1000 Fuente: guía de estudio oficial de Microsoft Learn

1. Prepare the data

25–30%

Get or connect to data

- Power BI Desktop usa Power Query para conectarse; el artículo de conectores indica qué capacidades admite cada origen, como DirectQuery. Para reutilizar un modelo semántico publicado se usa Obtener datos > Modelos semánticos de Power BI, que crea una conexión dinámica: exige permiso Build y no permite modificar el modelo, solo añadir medidas de informe y otros cálculos.
- En Power BI Desktop, Archivo > Opciones y configuración > Configuración de origen de datos permite editar permisos: cambiar credenciales, borrar permisos y fijar el nivel de privacidad, ya sea global o solo del archivo actual. Los niveles son Privado, Organizativo y Público; un origen Privado no se puede plegar en otros. Ignorar los niveles de privacidad puede exponer datos confidenciales.
- Import copia los datos en memoria y ofrece el conjunto de funciones más completo, pero requiere actualización. DirectQuery no importa datos y consulta el origen en cada interacción, útil con volúmenes enormes o datos casi en tiempo real. Direct Lake, en Fabric, carga tablas Delta de OneLake con el motor VertiPaq, y su actualización solo copia metadatos (framing).
- Un parámetro se crea desde Administrar parámetros o con Convertir en parámetro sobre una consulta de valor constante. Tiene nombre, tipo, valores sugeridos y valor actual. Se usa como argumento de un paso o de una función, y en la actualización incremental se exigen RangeStart y RangeEnd de tipo Fecha/Hora. En el servicio se pueden editar, pero no crear.

Profile and clean the data

- Las herramientas de generación de perfiles de datos del Editor de Power Query son Calidad de columna, Distribución de columna y Perfil de columna, y se activan en la pestaña Vista. Por defecto analizan las primeras 1000 filas, pero se pueden basar en todo el conjunto de datos. Calidad muestra valores válidos, con error y vacíos; la distribución indica valores distintos y únicos.
- Los errores de celda no impiden cargar la consulta y se pueden quitar, reemplazar por un valor o conservar para auditarlos. Los valores nulos se rellenan con Rellenar hacia abajo o hacia arriba, y Reemplazar valores corrige valores incorrectos o convierte celdas vacías en null. Con try ... otherwise se define un valor alternativo cuando una expresión produce error.
- Los errores de importación se diagnostican con el motivo, el mensaje y el detalle del error. Causas habituales: configuración regional que no coincide con el formato de origen, tipos inferidos solo de las primeras 200 filas, dimensiones incorrectas de una hoja de Excel, formatos antiguos que requieren Access Database Engine y errores Formula.Firewall por niveles de privacidad o pasos que combinan orígenes y referencias.

Transform and load the data

- El tipo de datos se define por columna y condiciona las transformaciones de Power Query. En orígenes no estructurados (Excel, CSV, texto) detecta tipos y encabezados inspeccionando los valores; una columna sin tipo explícito queda como Any y conviene evitarlo. Con Cambiar tipo usando configuración regional se interpretan bien fechas escritas en otro formato. El decimal fijo es siempre preciso, a diferencia del decimal de coma flotante.
- Las columnas nuevas se crean en la pestaña Agregar columna: personalizada con una fórmula M, condicional con cláusulas evaluadas de arriba abajo y una cláusula Else final, o a partir de ejemplos cuando conoces el resultado pero no la transformación. Las columnas condicionales nuevas no tienen tipo definido, así que hay que asignarlo. Una columna de índice empieza en 0 con incremento 1 y sirve de apoyo para otras transformaciones.
- Agrupar por resume filas según los valores de una o varias columnas; con la opción Avanzadas se eligen varias columnas de agrupación. Hay operaciones de columna (Suma, Promedio, Mediana, Mínimo, Máximo) y de fila (Contar filas, Contar filas distintas y Todas las filas). Todas las filas devuelve una tabla anidada por grupo, sobre la que se puede añadir una columna personalizada, por ejemplo con Table.Max, y expandir el registro resultante.
- Dinamizar convierte los valores de una columna en encabezados y agrega otra columna (por defecto con suma); No agregar espera un valor por intersección y da error si hay varios. Anular dinamización convierte columnas en pares Atributo y Valor; Anular dinamización de otras columnas incorpora las columnas nuevas del origen al actualizar. Transponer gira la tabla y pierde los encabezados, que se recuperan con Usar primera fila como encabezados.
- El conector JSON aplana los datos con detección automática de tablas; para varios archivos se usa un conector de carpeta. Analizar (Parse) interpreta una columna de texto como JSON, que da valores Record, o como XML, que da valores Table; en la pestaña Transformar modifica la columna y en Agregar columna crea otra. Expandir una lista crea una fila por valor repitiendo las demás columnas, y expandir un registro crea una columna por campo.

2. Model the data

25–30%

Design and implement a data model

- Las propiedades de columna definen cómo trata Power BI cada campo en los objetos visuales. La categoría de datos aclara el significado de un valor (por ejemplo, geográfico) para que se represente bien en un mapa. Ordenar por columna exige que ambas columnas tengan el mismo nivel de detalle y que la columna de orden tenga un valor único por cada valor ordenado, como el nombre del mes ordenado por su número.
- Una dimensión de rol múltiple, como la fecha de pedido, envío y entrega, filtra los mismos hechos de formas distintas. En Power BI solo puede haber una relación activa entre dos tablas; las demás son inactivas y solo se usan con USERELATIONSHIP. Para filtrar por varios roles a la vez, se crea una tabla por rol, por ejemplo con una tabla calculada que copie la original, cada una con su relación activa.
- La cardinalidad puede ser varios a uno (la más habitual y predeterminada), uno a uno, uno a varios o varios a varios, y requiere que al menos un lado tenga valores únicos salvo en varios a varios. La dirección del filtro cruzado es única de forma predeterminada; Ambas trata las tablas como una sola, pero puede crear ambigüedad y afectar al rendimiento, por lo que conviene minimizarla y preferir CROSSFILTER en una medida.
- Una tabla de fechas tiene una fila por día y una columna de fecha única, sin vacíos ni fechas que falten, que abarque años completos. Puede cargarse desde un almacén, generarse en Power Query o crearse con CALENDAR y CALENDARAUTO. Marcarla como tabla de fechas valida esas reglas y es obligatorio si la relación usa una clave entera yyyymmdd. La fecha y hora automáticas crean una tabla oculta por columna.
- Las columnas y tablas calculadas usan DAX sobre datos ya cargados, se calculan al actualizar y se guardan en el modelo, por lo que no cambian con la interacción del usuario y aumentan su tamaño. Una columna calculada sirve para segmentaciones, filtros, filas y columnas; una tabla calculada, para datos intermedios como una UNION o un CROSS JOIN. Para valores que responden a los filtros del informe se usa una medida.

Create model calculations by using DAX

- Una medida es un cálculo DAX que se evalúa bajo demanda y cambia con los filtros y segmentaciones del informe; su resultado no se guarda en el modelo. Las agregaciones simples usan SUM o AVERAGE sobre una columna, y SUMX o AVERAGEX cuando hay que evaluar una expresión fila a fila. AVERAGE ignora las celdas vacías, pero cuenta los ceros.
- CALCULATE evalúa una expresión en un contexto de filtro modificado. Sus filtros pueden ser expresiones booleanas, expresiones de tabla o modificadores como REMOVEFILTERS, ALL o KEEPFILTERS; un filtro sobre una columna ya filtrada sustituye al existente, salvo que se use KEEPFILTERS. Sin filtros, convierte el contexto de fila en contexto de filtro.
- Las funciones de inteligencia de tiempo de DAX permiten comparar y acumular valores por días, meses, trimestres y años, con familias como DATESYTD y TOTALYTD, SAMEPERIODLASTYEAR, DATEADD y PARALLELPERIOD. Antes de usarlas hay que marcar como tabla de fechas una tabla que contenga una columna de fechas.
- DAX ofrece funciones estadísticas como MEDIAN y MEDIANX (mediana de una columna o de una expresión por fila), PERCENTILE.INC (k entre 0 y 1), RANKX, DISTINCTCOUNT y las desviaciones STDEV.S para muestras y STDEV.P para la población completa. Los recuentos distintos no son aditivos en los totales.
- Las medidas semiaditivas, como los saldos o las existencias, no se agregan igual en todas las dimensiones: en el tiempo se toma un valor puntual, no la suma. CLOSINGBALANCEMONTH, QUARTER y YEAR evalúan la expresión en la última fecha del periodo, y LASTNONBLANK devuelve el último valor con datos en el contexto actual.

Optimize model performance

- Conserva solo las columnas que sirven para informes (filtrar, agrupar, resumir) o para la estructura del modelo (relaciones, cálculos, roles de seguridad); el resto se puede quitar (filtrado vertical). Carga también pocas filas, filtrando por tiempo (p. ej. historial relativo con parámetros) o por entidad (filtrado horizontal). Un modelo más pequeño se actualiza antes y calcula más rápido.
- Performance Analyzer registra cuánto tarda cada visual y lo desglosa en consulta DAX, DirectQuery, visualización y otros; los resultados se pueden exportar a JSON. Desde ahí se copia la consulta DAX de un visual o se ejecuta en la vista de consulta DAX, donde se inspeccionan y prueban medidas con DEFINE MEASURE antes de actualizar el modelo.
- Cargar datos ya resumidos (agrupar por fecha, cliente y producto, o por fecha a nivel de mes) eleva el grano de la tabla de hechos y es una de las técnicas más eficaces para reducir el modelo. A cambio se pierde detalle: ya no se informa por día o por línea de pedido. Se puede compensar con tablas en DirectQuery para el detalle en una página de drillthrough.

3. Visualize and analyze the data

25–30%

Create reports

- Elige el visual según los datos y el objetivo: columnas y barras para comparar categorías, líneas y áreas para tendencias en el tiempo, gráficos combinados para medidas con escalas distintas, anillo y treemap para partes de un todo, dispersión para relaciones entre dos valores, tablas y matrices para valores exactos, y tarjetas, KPI y medidores para una cifra clave. Considera también la audiencia y el espacio disponible.
- El panel de formato del visual (icono del pincel) ofrece opciones que dependen del tipo de visual: colores, ejes, etiquetas de datos, títulos, fondos y bordes. Puedes restablecer un visual entero o una sección a los valores predeterminados, aplicar estilos predefinidos a tablas y matrices y usar fx para el formato condicional. El formato aplicado a un visual prevalece sobre el tema.
- El visual de narrativa de Copilot resume todo el informe, páginas concretas o visuales seleccionados, y lee lo que hay en el lienzo, no el modelo semántico, por lo que conviene dar nombres claros a visuales y ejes. El cuerpo no se edita directamente: se ajusta con mensajes, y las notas al pie resaltan los visuales de origen. No se actualiza solo al filtrar; el lector usa el botón de actualizar.
- Un tema aplica colores y formato predeterminado a todos los visuales: se elige uno integrado o de la organización en la galería de temas, o se importa, personaliza y descarga un archivo JSON. Conviene usar colores del tema para que los visuales se actualicen al cambiarlo; el formato de un visual concreto se mantiene hasta restablecerlo, y aplicar un tema integrado o importado elimina las personalizaciones previas.
- El formato condicional cambia colores, barras de datos, iconos o vínculos según los valores, con tres estilos: degradado (numérico), reglas y valor del campo (nombres de color o códigos hexadecimales, normalmente de una medida). En tablas y matrices se aplica a valores, totales y subtotales, no a encabezados. Con porcentajes se escriben decimales y Número, y los valores NaN impiden degradados y reglas porcentuales automáticos; usa DIVIDE.

Enhance reports for usability and storytelling

- Un marcador guarda el estado de una página: filtros y segmentaciones, selecciones de resaltado cruzado, orden, nivel de exploración y visibilidad de objetos del panel Selección. Las opciones Datos, Mostrar y Página actual controlan qué aplica, y los marcadores se pueden agrupar, enlazar a botones o mostrar con un navegador. Si cambias la visibilidad después de guardar, debes actualizar el marcador.
- Un tooltip de informe es una página con el tamaño de página Tooltip y el interruptor Tooltip activado en Información de página. Los campos que arrastres a Campos del tooltip hacen que aparezca en los visuales que los usan, y también se puede asignar a mano o volver a Automático. No son interactivos y los lectores de pantalla no los leen, así que no los uses para información esencial.
- De forma predeterminada, los visuales de una página se filtran y resaltan entre sí. Con Editar interacciones, el autor elige en cada visual de destino el icono Filtro, Resaltar o Ninguno, y esos cambios se guardan con el informe. La exploración en profundidad no afecta a otros visuales salvo que se configure para toda la página, y los lectores en vista de lectura no pueden cambiar las interacciones.
- El navegador de páginas se sincroniza solo con las páginas del informe: etiquetas, orden y altas, bajas y cambios de nombre. El navegador de marcadores muestra todos los marcadores o solo un grupo, y permite anular la selección con un marcador de destino. Los botones, formas e imágenes admiten acciones como Atrás, Navegación de páginas, Marcador o Dirección URL web, y el destino puede depender de una selección con formato condicional.
- Ordenar por columna hace que una columna defina el orden de otra, por ejemplo el número de mes para los nombres de mes; ambas deben tener la misma granularidad y la de ordenación no puede tener valores duplicados. En los visuales, el menú Más opciones ofrece ordenar ascendente, descendente y por eje, y Mayús permite ordenar por varias columnas. Algunos visuales, como tarjetas o medidores, no se pueden ordenar.

Identify patterns and trends

- Con clic derecho en un punto de datos, Analyze ofrece Explain the increase o decrease y Find where this distribution is different. Explain compara el punto con el anterior y muestra qué categorías cambiaron más su contribución; el resultado puede verse como cascada, dispersión, columnas 100 % apiladas o cinta, y en edición se añade al informe. No está disponible para el primer punto de datos.
- Agrupar (Ctrl+clic y Group) reúne elementos de un visual en una categoría nueva, editable con Edit groups. Binning define un tamaño de intervalo para campos numéricos o de fecha, también en columnas calculadas pero no en medidas. Clustering, con Automatically find clusters en un gráfico de dispersión, crea un campo categórico con grupos de puntos similares y admite fijar el número de clústeres.
- Influenciadores clave analiza una métrica con campos en Explain by y ordena los factores que la determinan; Top segments muestra combinaciones de factores y Expand by fija el nivel de análisis de una medida. El árbol de descomposición usa una medida en Analyze y permite profundizar por dimensiones en cualquier orden, con divisiones de IA de valor alto o bajo y niveles bloqueables. El visual Narrative resume visuales en texto.
- El panel Analytics añade líneas de referencia (constante, mínimo, máximo, promedio, mediana, percentil), barras de error, previsión y anomalías. Las barras de error se definen por campo, con límites absolutos o relativos, o por porcentaje. La previsión requiere datos de tiempo, solo está disponible en gráficos de líneas y admite ajustes como la longitud de la previsión y el intervalo de confianza.
- La detección de anomalías se activa con Anomalies en el panel Analytics de un gráfico de líneas con serie temporal en el eje, sin leyenda ni varios valores. Marca los puntos fuera del rango esperado, la sensibilidad controla cuántos se marcan y Explain by limita los campos usados para explicarlos. No funciona junto con líneas de previsión, mínimo, máximo, promedio, mediana o percentil.

4. Manage and secure Power BI

15–20%

Create and manage workspaces and assets

- Las áreas de trabajo se crean en el servicio Power BI, no en Power BI Desktop, y cada una necesita un nombre único. Entre los ajustes opcionales están la lista de contactos, el OneDrive del área de trabajo y permitir que los Contributor actualicen la aplicación. El acceso se concede con los roles Admin, Member, Contributor y Viewer, y solo un Admin puede eliminar el área de trabajo.
- Una aplicación empaqueta contenido de un área de trabajo con navegación integrada, y el área de trabajo es el espacio de preparación: los cambios llegan a los usuarios solo al actualizar la aplicación. Se pueden definir varias audiencias que ven contenido distinto y hay que incluir los informes de origen de un panel para que sus mosaicos funcionen. Los usuarios no pueden modificar el contenido de la aplicación.
- Publicar desde Power BI Desktop envía el modelo semántico y sus informes a un área de trabajo, igual que cargar un archivo .pbix con Obtener datos. Al volver a publicar se reemplaza el modelo semántico existente, se pueden romper visuales si se renombran o eliminan campos, y no se puede pasar del modo de importación a DirectQuery. Publicar exige un rol que permita crear contenido en el área de destino.
- Un panel es una sola página de mosaicos que se crea en el servicio Power BI, no en Desktop, anclando visuales de informes o una página completa como mosaico dinámico. Puede combinar informes y modelos semánticos distintos y admite mosaicos de texto, imagen, vídeo, contenido web y streaming. Los paneles no admiten filtros ni segmentaciones, y las alertas de datos solo se aplican a tarjetas, medidores y KPI.
- Se usa un área de trabajo para colaborar en equipo, una aplicación para distribuir contenido seleccionado de solo lectura a audiencias amplias, y los vínculos o el acceso directo para personas concretas. Los tipos de vínculo son Personas de su organización, Personas concretas y Personas con acceso existente. Compartir un informe comparte también su modelo semántico, así que conviene restringir los datos con seguridad a nivel de fila.

Secure and govern Power BI items

- Un área de trabajo tiene cuatro roles: Admin, Member, Contributor y Viewer, que se pueden asignar a personas o grupos. Si alguien está en varios grupos, obtiene el rol más alto. Admin gestiona el área y a sus usuarios, Member añade personas con roles inferiores y publica apps, Contributor crea y edita contenido y Viewer solo consulta.
- Además de los roles del área de trabajo, se puede compartir un elemento concreto para dar permisos solo sobre él a quien no tiene rol. Al compartir siempre se concede al menos lectura. Los vínculos pueden ser para la organización, para personas específicas o para quien ya tiene acceso, y los permisos heredados de un rol no se cambian en la página de permisos del elemento.
- El acceso a un modelo semántico se controla con los permisos Read, Build, Reshare y Write. Read permite usar los informes existentes, Build permite crear contenido nuevo y usar Analizar en Excel, Reshare permite dar acceso a otros y Write permite modificar y volver a publicar. Los roles del área los conceden implícitamente y el resto se gestiona al compartir o en la página de permisos.
- Los roles de RLS se definen en Power BI Desktop con filtros DAX que devuelven verdadero o falso por fila; pueden ser estáticos o dinámicos con USERPRINCIPALNAME. Se publican con el modelo y se validan con Probar como rol. RLS solo restringe a los Viewer del área de trabajo, no a Admin, Member ni Contributor.
- Los miembros de un rol de RLS se asignan en el servicio Power BI, desde Seguridad en el modelo semántico, no en Desktop. Se admiten usuarios y grupos de seguridad de Entra, grupos habilitados para correo y grupos de distribución, pero no grupos de Microsoft 365 ni entidades de servicio. Los roles son aditivos y a los invitados externos conviene añadirlos directamente por correo.