Promts GPT para precariedad.  
  
**✅ Preguntas clave para planificar el trabajo**

**📌 Sobre el Alcance y Objetivos:**

1. ¿Qué nivel de detalle esperan en el informe (ejecutivo o técnico)?
2. ¿Desean realizar análisis solo a nivel nacional o desagregar por región?
3. ¿Incluirán comparaciones con literatura académica o solo resultados empíricos?
4. ¿Habrá un anexo metodológico detallado (fuentes, cálculos, códigos)?

**📂 Sobre la Organización del Proyecto:**

1. ¿Cómo quieren estructurar las carpetas y scripts en GitHub?
2. ¿Definiremos convenciones para nombres de scripts, funciones y objetos?
3. ¿Qué formato de exportación prefieren para tablas y gráficos (.csv, .xlsx, .png, .html)?
4. ¿Quién revisará los *pull requests* y cómo gestionarán las revisiones?

**📊 Sobre el Análisis de Precariedad:**

1. ¿Cómo definirán exactamente los signos de precariedad (códigos de la EPH específicos)?
2. ¿Qué criterio usarán para el cruce de precariedad con sectores y rangos etarios?
3. ¿Calcularán tasas ajustadas o solo tasas simples?
4. ¿Harán test de significancia para las diferencias interanuales?

**💰 Sobre el Monitor de Ingresos:**

1. ¿Qué índice de precios usarán para ajustar los ingresos reales (IPC, CBA, CBT)?
2. ¿Cómo manejarán los valores extremos de ingresos (*outliers*)?
3. ¿Desagregarán ingresos por deciles o solo promedios?
4. ¿Quieren mostrar *gap* salarial entre hombres y mujeres?

**📈 Sobre la Visualización y Dashboards:**

1. ¿Qué tipo de gráficos serán prioritarios (líneas, barras, mapas, pirámides)?
2. ¿Requieren comparaciones trimestrales o solo interanuales?
3. ¿Incluirán filtros interactivos por sexo, edad, sector?
4. ¿Prefieren *dashboards* estáticos (flexdashboard) o interactivos (shiny)?

**🛠️ Sobre la Programación:**

1. ¿Quieren que cree funciones reutilizables para tasas y gráficos?
2. ¿Automatizamos reportes trimestrales con rmarkdown::render()?
3. ¿Desean que prepare scripts de validación de datos (duplicados, faltantes, inconsistencias)?
4. ¿Usarán librerías específicas o dejamos tidyverse como base?

**🚀 Próximo paso:**

Si respondes estas preguntas o me aclaras qué puntos priorizar, puedo:

1. **Crear un esquema del .rmd** con la estructura y los bloques de análisis.
2. **Escribir las primeras funciones R** para calcular tasas y signos de precariedad.
3. **Proponer ejemplos de visualización** según los indicadores definidos.

¿Con cuál quieres que arranquemos primero? 😊

1. Técnico 2. nacional 3. Empírico con mencion a precariedad ceped 4. sí, con anexo metodológico 5. no lo tengo claro 6. si 7. .htpml 8. ambos integrantes revisan y gestionan 9. se define de la siguiente manera filter(ESTADO == 1, CAT\_OCUP == 3) %>% # Ocupados asalariados mutate( nivel.ed = factor( case\_when(NIVEL\_ED %in% c(7,1,2,3) ~ "Menor a Secundaria", NIVEL\_ED %in% c(4,5) ~ "Secundaria Completa", NIVEL\_ED == 6 ~ "Superior Completo", TRUE ~ "Ns/Nr"), levels = c("Menor a Secundaria","Secundaria Completa","Superior Completo")), tamanio.establec = factor( case\_when(PP04C %in% 1:6 ~ "Pequeño", PP04C %in% 7:8 ~ "Mediano", PP04C %in% 9:12 ~ "Grande", PP04C %in% 99 ~ "Ns/Nr"), levels = c("Pequeño","Mediano","Grande","Ns/Nr")), descuento\_jubil = case\_when(PP07H == 1 ~ "Si", PP07H == 2 ~ "No"), part.time.inv = case\_when(PP3E\_TOT < 35 & PP03G == 1 ~ "Si", TRUE ~ "No"), tiempo.determinado = case\_when(PP07C == 1 ~ "Si", TRUE ~ "No")) 10. al menos un signo de precariedad 11. simples 12. si 13. IPC 14. se excluye pero no es claro como se define el límite de corte 15. promedios por rango etario y sexo, deciles para distribucion de totales sin desagregar por rangos etarios. 16. solo para el caso según grado de precariedad 17. barras, mapas de calor y pirámides 18. solo interanuales 19. si, y buenas etiquetas 20. estáticos 21. si 22. si 23. por ahora no- 24. se utiliza mucho la librería eph Teniendo esto en cuenta, redacta una serie de pasos a seguir y preguntas que puedan ayudar a mejorar los prompts en cada paso.  
  
Aquí tienes una guía de pasos clave para tu proyecto, junto con preguntas que pueden ayudarte a mejorar el enfoque en cada etapa:

**📂 1. Organización del Proyecto**

* **Pasos:**
  + Crear una estructura clara de carpetas (data/, scripts/, outputs/, plots/, etc.).
  + Establecer una convención de nombres (snake\_case) y roles en GitHub (revisiones mediante *pull requests*).
* **Preguntas:**
  + ¿Ya tienen un archivo README.md que explique la estructura del proyecto?
  + ¿Definieron una convención para los nombres de ramas y mensajes de *commit*?

**📊 2. Carga y Preparación de Datos**

* **Pasos:**
  + Automatizar la descarga y carga de bases con eph::get\_microdata().
  + Crear un script para limpieza y unificación de trimestres.
* **Preguntas:**
  + ¿Cómo tratarán las ponderaciones (PONDERA, PONDIIO)?
  + ¿Cómo documentarán las transformaciones aplicadas (scripts o anexo metodológico)?

**📈 3. Definición de Signos de Precariedad**

* **Pasos:**
  + Reunir las definiciones en una función (calcular\_precariedad()), usando su código de mutate().
  + Añadir validaciones para evitar errores de codificación (e.g., PP07H solo toma valores 1 o 2).
* **Preguntas:**
  + ¿Cómo diferenciarán asalariados privados y públicos?
  + ¿Usarán etiquetas para cada signo o solo una variable binaria (signos\_precariedad)?

**📊 4. Cálculo de Tasas y Tablas**

* **Pasos:**
  + Definir funciones (calcular\_tasa(), comparar\_interanual()).
  + Generar tablas de frecuencia (tabyl() de janitor).
* **Preguntas:**
  + ¿Qué método usarán para diferencias interanuales (porcentaje o puntos porcentuales)?
  + ¿Cómo manejarán celdas con pocos casos (ocultar, agrupar)?

**💰 5. Análisis de Ingresos Asalariados**

* **Pasos:**
  + Crear una función calcular\_ingreso\_real() ajustando por IPC.
  + Calcular promedios y deciles según precariedad y grupo poblacional.
* **Preguntas:**
  + ¿Ya tienen la serie de IPC para el ajuste?
  + ¿Qué harán con ingresos fuera de rango (extremos, negativos)?

**📊 6. Visualización (Dashboard y Gráficos)**

* **Pasos:**
  + Usar ggplot2 para gráficos estáticos (barras, pirámides, mapas de calor).
  + Asegurar títulos claros y escalas comparables.
* **Preguntas:**
  + ¿Qué colores usarán para distinguir sexo, edad, y precariedad?
  + ¿Cómo planean mostrar la evolución interanual (líneas, facetas)?

**📑 7. Documentación y Anexo Metodológico**

* **Pasos:**
  + Añadir explicaciones en RMarkdown con secciones (## Metodología, ## Resultados).
  + Guardar el anexo metodológico en anexo.html desde el .rmd.
* **Preguntas:**
  + ¿Cómo documentarán las decisiones sobre *outliers* y ponderaciones?
  + ¿Incluirán ejemplos de código en el anexo?

Recomendaciones Notebook LM  
  
Para abordar el análisis de la precariedad laboral en Argentina utilizando la Encuesta Permanente de Hogares (EPH) y considerando el trabajo final en .rmd con GitHub, aquí hay algunas recomendaciones y consideraciones clave, organizadas en función de los aspectos que mencionaste:

**Flujo de trabajo con GitHub y R**

1.**Estructura del proyecto**:

◦Utiliza un repositorio principal (main project) en GitHub y ramas (branches) para cada tarea o análisis específico1.

◦Divide el trabajo en funciones lógicas y scripts separados para facilitar la colaboración y el mantenimiento2.

2.**Colaboración**:

◦Establece un flujo de trabajo claro con tu compañero, utilizando pull requests para la revisión del código antes de fusionar los cambios a la rama principal.

◦Utiliza herramientas de gestión de proyectos en GitHub (issues, project boards) para coordinar tareas y prioridades.

3. **Optimización de funciones**:

◦Escribe funciones reutilizables para tareas comunes (ej. limpieza de datos, cálculos de tasas)2.

◦Utiliza la vectorización en R para mejorar el rendimiento de las operaciones2.

**Organización de Scripts**

1.**Script principal**:

◦Un script central (main.R o similar) que coordine la ejecución de los demás scripts.

◦Este script puede cargar los datos, llamar a las funciones de limpieza y análisis, y generar los outputs2.

2.**Scripts de funciones**:

◦Scripts separados para cada conjunto de funciones relacionadas (ej. cleaning\_functions.R, analysis\_functions.R).

◦Documenta cada función con comentarios claros sobre su propósito, argumentos y valores de retorno.

3 **Script de visualización**:

◦Un script dedicado a la generación de gráficos y tablas2.

◦ Utiliza un estilo consistente para las visualizaciones y guarda los gráficos en formatos adecuados (ej. PNG, PDF).

**Variables de la EPH a Utilizar**

Aquí hay una lista de variables de la EPH que pueden ser relevantes para el análisis de la precariedad laboral, organizadas por tema:

•

**Identificación y Demografía**:

◦CODUSU: Código para distinguir viviendas, permite aparear hogares y personas3.

◦NRO\_HOGAR: Código para distinguir hogares, permite aparearlos con personas4....

◦COMPONENTE: Número de miembro del hogar6.

◦ANO4: Año de relevamiento4....

◦TRIMESTRE: Número de trimestre4....

◦CH06: Edad6.

◦CH04: Sexo6.

◦NIVEL\_ED: Nivel educativo alcanzado7....

•**Condición de Actividad**:

◦ESTADO: Condición de actividad (ocupado, desocupado, inactivo, menor de 10 años)7....

◦CAT\_OCUP: Categoría ocupacional (patrón, cuenta propia, obrero/empleado, trabajador familiar)7....

◦CAT\_INAC: Categoría de inactividad (jubilado, rentista, estudiante, ama de casa, etc.)9....

•**Búsqueda de Empleo (Desocupados)**:

◦PP02C1 a PP02C8: Formas de búsqueda de trabajo (contactos, currículum, establecimientos, etc.)9....

◦PP02E: Razón por la que no buscó trabajo en los últimos 30 días16....

◦PP02H: Buscó trabajo en los últimos 12 meses16....

◦PP02I: Trabajó en los últimos 12 meses17....

◦PP10A: Tiempo buscando trabajo23....

•**Características del Empleo (Ocupados)**:

◦PP03C: ¿Tenía un solo empleo/ocupación/actividad?21....

◦PP03D: Cantidad de ocupaciones21....

◦PP3E\_TOT: Total de horas trabajadas en la ocupación principal2....

◦PP3F\_TOT: Total de horas trabajadas en otras ocupaciones21....

◦PP03G: ¿Quería trabajar más horas?2....

◦PP03H: Si hubiera conseguido más horas, ¿podía trabajarlas?22....

◦PP03I: Buscó trabajar más horas en los últimos 30 días22....

◦PP03J: Buscó otro empleo/ocupación/actividad22....

◦INTENSI: Intensidad de trabajo32...

◦PP04A: Tipo de empleador (estatal, privado, otro)34....

◦PP04B\_COD: Actividad del negocio/empresa/institución (CAES-Mercosur)35....

◦PP04C: Cantidad de personas que trabajan en el establecimiento2....

◦PP04G: Dónde realiza principalmente sus tareas38....

◦PP05H: Tiempo trabajando en ese empleo de forma continua47....

◦PP07A: Tiempo trabajando en ese empleo en forma continua (asalariados)51....

◦PP07C: ¿Ese empleo tiene tiempo de finalización? (asalariados)2....

◦PP07D: ¿Por cuánto tiempo es ese trabajo? (asalariados con tiempo de finalización)56....

◦PP07H: ¿Tiene descuento jubilatorio? (asalariados)2....

◦PP07I: ¿Aporta por sí mismo a algún sistema jubilatorio?62....

◦PP07K: ¿Le dan recibo de sueldo?63....

•**Ingresos**:

◦PP08D1: Monto total de sueldos/jornales, salario familiar, horas extras, etc.66....

◦P21: Monto de ingreso de la ocupación principal70....

•**Desempleo anterior**:

◦PP11A: Tipo de negocio en el que trabajaba27....

◦PP11O: Razón principal por la que dejó ese trabajo79....

◦PP11N: ¿Le hacían descuento jubilatorio?80....

•**Variables del hogar**:

◦V8: Ingresos de alquileres86.

◦V9: Ingresos de negocios en los que no trabajan86.

◦V10: Ingresos de intereses o rentas por plazos fijos/inversiones86.

**Recomendaciones para el Análisis**

1.**Definición de Precariedad**:

◦Define claramente qué características consideras como "signos de precariedad"87.

◦Considera la informalidad (empleo en negro, falta de aportes jubilatorios), inestabilidad (contratos a corto plazo, falta de indemnización por despido), bajos ingresos, y falta de acceso a beneficios sociales (obra social, vacaciones pagas)1....

2. **Categorización del Mercado Laboral**:

◦Crea variables dummy para identificar a los trabajadores con diferentes grados de precariedad (0, 1, 2 o 3 signos)87....

◦Analiza la distribución de los trabajadores según estas categorías por sexo, edad y nivel educativo87....

3. **Análisis de la Evolución Temporal**:

◦Compara los trimestres 3ro y 4to de 2022 contra 2023, y 1ro y 2do de 2023 contra 202489.

◦Calcula tasas de variación interanuales para los diferentes grupos poblacionales87....

4 **Visualización de Datos**:

◦Utiliza gráficos de barras para comparar la cantidad de casos y el peso relativo de los diferentes grupos poblacionales2.

◦Emplea gráficos de líneas para mostrar la evolución de las tasas de variación a lo largo del tiempo2.

◦Crea tablas para presentar los resultados de forma clara y concisa2.

◦Considera el uso de *dashboards* para una visualización interactiva de los resultados2.

**Ejemplos de Visualización**

1. **Gráfico de Barras**:

◦Cantidad de asalariados precarios según rango etario y sexo, comparando 2022 y 2023.

◦Porcentaje de trabajadores informales según nivel educativo.

2**Gráfico de Líneas**:

◦Evolución de la tasa de empleo no registrado por grupo de edad.

◦Variación interanual de los ingresos de los asalariados según grado de precariedad.

3**Tabla**:

◦Distribución de los ocupados por categoría ocupacional (asalariado, cuenta propia, etc.) y nivel educativo.

◦Ingresos promedio de los asalariados según sector de actividad y grado de precariedad.

**Cuestiones de Programación**

1 **Limpieza de Datos**:

◦Gestiona los valores faltantes (NA) de forma adecuada (imputación, eliminación)90....

◦Recodifica las variables categóricas a factores con etiquetas descriptivas2.

◦Verifica la consistencia de los datos (ej. rangos de edad, valores de ingresos)91.

2 **Ponderación**:

◦Utiliza la variable PONDERA para ponderar los resultados y obtener estimaciones representativas de la población4....

◦Considera el uso de PONDII para el tratamiento del ingreso total individual y PONDIH para el ingreso total familiar, que corrigen por la no respuesta de ingresos93.

3 **Análisis de Ingresos**:

◦Ten en cuenta los códigos especiales para los montos de ingreso: -7 (No tenía esa ocupación), -8 (No tuvo ingresos), -9 (No responde)94....

◦Utiliza los campos P21 (ingreso de la ocupación principal) y TOT\_P12 (ingreso de otras ocupaciones) para analizar los ingresos laborales72....

4.**Clasificadores**:

◦ Utiliza el Clasificador de Actividades Económicas para Encuestas Sociodemográficas del Mercosur (CAES-Mercosur) para analizar los sectores de actividad36....

◦ Considera el Clasificador Nacional de Ocupaciones (CNO) para analizar las ocupaciones38....

**Recomendaciones Adicionales**

• **Documentación del INDEC**:

◦ Consulta los cuestionarios de la EPH y los documentos metodológicos del INDEC para comprender el diseño de la encuesta y las definiciones de las variables6....

◦ Accede a la página web del INDEC para obtener información actualizada sobre la EPH y otros documentos relevantes6....

• **Errores de Muestreo**:

◦ Consulta las tablas de errores de muestreo publicadas por el INDEC para dimensionar el error muestral asociado a las estimaciones10 5....

• **Tamaño de la Muestra**:

◦ Ten en cuenta que un tamaño de muestra pequeño puede resultar insuficiente para el cálculo de algunas estimaciones, especialmente en subpoblaciones 105....

◦ Considera utilizar la base general o agrupamientos de aglomerados para incrementar el número de casos muestrales 105....