La **hipótesis central** de este análisis se centra en evaluar la relación entre el consumo de alcohol y el rendimiento académico de los estudiantes. El interés particular de esta primera fase de análisis consiste en determinar cómo las prácticas de consumo de alcohol, incluyendo la frecuencia de socialización con alcohol, la cantidad de bebidas consumidas por salida y las clases perdidas debido a razones relacionadas con el alcohol, pueden influir en los promedios académicos de los estudiantes. Se busca explorar si un mayor consumo de alcohol está asociado con una disminución en el rendimiento académico, manifestado a través de calificaciones más bajas o un mayor número de materias reprobadas. Esta hipótesis se basa en la premisa de que el consumo excesivo de alcohol podría afectar negativamente la capacidad de los estudiantes para mantener un alto nivel de rendimiento académico, debido a factores como la fatiga, la falta de concentración y la pérdida de tiempo de estudio efectivo.

LA BASE DE DATOS CONSISTE EN UNA ENCUESTA A ESTUDIANTES SOBRE SU CONSUMO DE ALCOHOL Y OTRAS VARIABLES DEMOGRÁFICAS.

**Información Demográfica**: Incluye datos como género y año académico del estudiante.

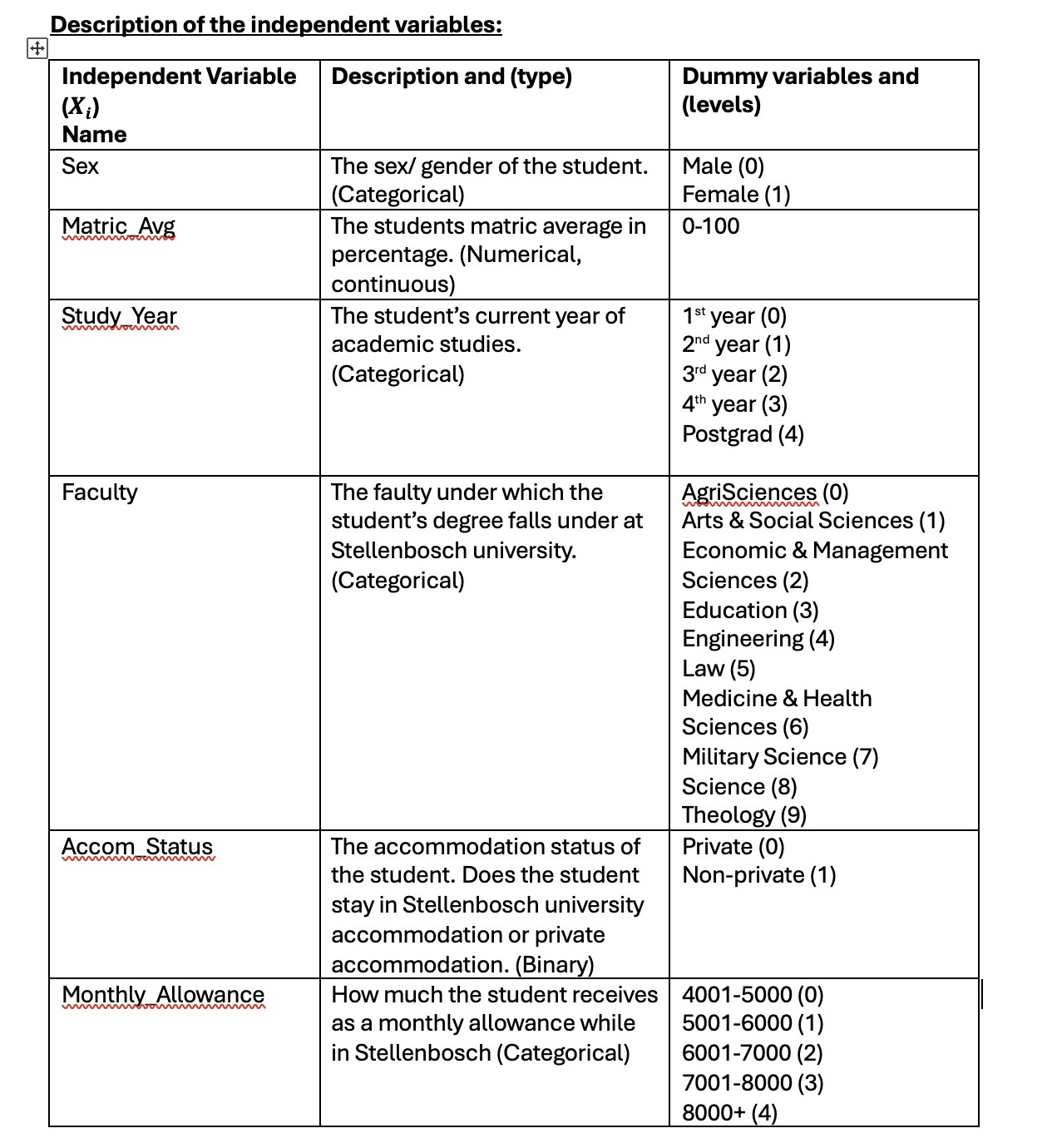
**Rendimiento Académico**: Se recoge tanto el promedio de notas del último año de secundaria (bachillerato) como el promedio académico del año 2023.

**Facultad**: Indica a qué facultad pertenece el grado que está cursando el estudiante.

**Estilo de Vida y Hábitos Personales**:

* + **Alojamiento**: Estado de alojamiento del estudiante durante el año 2023.
  + **Asignación Mensual**: Información sobre la cantidad de dinero que recibe el estudiante cada mes.
  + **Beca o Subsidio**: Si el estudiante estaba recibiendo algún tipo de beca o subsidio durante el 2023.
  + **Horas de estudio adicional**: Cantidad de horas de estudio adicional por semana.
  + **Socialización y Consumo de Alcohol**: Frecuencia de salidas a socializar o fiestas durante la semana, cantidad de bebidas alcohólicas consumidas durante estas salidas, y número de clases perdidas por razones relacionadas con el alcohol.
  + **Relación Amorosa**: Si el estudiante tiene una relación amorosa actualmente.
  + **Aprobación de los Padres sobre el Consumo de Alcohol** y **Relación con los Padres**: Explora la dinámica familiar respecto al consumo de alcohol y la fortaleza de la relación con los padres.
  + **Rendimiento Académico Adverso**: Incluye variables como el número de módulos académicos reprobados.

Descripción de variables independientes presentes en la base:



Procedimiento:

1. El primer gráfico realizado en el análisis consiste en un mapa de calor para poder evaluar visualmente la calidad de los datos, y la presencia de NaN. Dado que el número de datos faltantes en las variables a analizar es bajo, no es necesario realizar una imputación de datos.
2. Luego procedo a realizar un análisis exploratorio de los datos, mediante diversos gráficos (realizados en matplotlib y seaborn), para ello comienzo graficando la cantidad de estudiantes que hay por año de cursada, la cantidad de clases perdidas por consumo de alcohol, la conformación de la sexo de la muestra, la mensualidad de los estudiantes y la distribución de la muestra de acuerdo al promedio académico. Asimismo gráfico el promedio académico para mujeres y para varones, para descartar que de antemano existan diferencias significativas en el rendimiento por sexo.
3. En esta etapa exploraré la hipótesis secundaria, que **el consumo de alcohol es mayor en hombres que en mujeres.** La hipótesis es confirmada como puede observarse en el heatmap graficado, una mayor cantidad de varones reportan consumir 8 bebidas alcohólicas o mas por noche.
4. Para evaluar si aquellas **personas con mejor rendimiento académico consumen menos bebidas alcohólicas,** realizo un gráfico de frecuencias, buscando observar gráficamente alguna tendencia. No hay tendencia distinguible, tanto individuos con buen promedio como con bajo promedio se distribuyen uniformemente a lo largo de las variables de consumo de alcohol.
5. En esta instancia **busqué evaluar si el patrón de consumo de alcohol se relaciona con el apoyo familiar**, medido como cercanía a los padres. No se observa ningún patrón definitivo en el heatmap.
6. Mediante gráfico de boxplot se **buscó evaluar si la cantidad de veces que un individuo sale por semana impacata en el rendimiento académico**, nuevamente no se observó ninguna tendencia puntual.
7. Finalmente mediante gráfico de barras apiladas se buscó **evaluar gráficamente si la cantidad de bebidas consumidas por noche impactaba en el promedio**. No observándose ninguna tendencia.

**RESULTA IMPORTANTE DESTACAR QUE TODO ANÁLISIS REALIZADO PARA CONTRASTAR LAS HIPOTESIS PLANTEADAS FUE GRÁFICO, NO SE REALIZARON ANÁLISIS ESTADÍSITICOS, QUE SERÍAN CONCLUYENTES PARA CONFIRMAR O REFUTAR HIPOTESIS. EN UNA SEGUNDA ETAPA PUEDEN REALIZARSE CORRELACIONES BIVARIADAS, ANOVA O TEST T PARA BUSCAR DIFERENCIAS SIGNIFICATIVAS ENTRE LAS VARIABLES EXPLORADAS.**