



PARCIAL I: ANÁLISIS EXPLORATORIO DE DATOS DEL CONJUNTO DE DATOS VAST CHALLENGE 2022

Alumno: Nelson Jorge Apaza Apaza

ANÁLISIS EXPLORATORIO DE DATOS DEL CONJUNTO DE DATOS VAST CHALLENGE 2022

El conjunto de datos **VAST Challenge 2022** consiste en **datos urbanos** ficticios que **emulan tareas y problemas de análisis visual** del mundo real [He et al.]. Estos retos, incluyendo ediciones anteriores, también han sido utilizados en **proyectos de analítica inmersiva colaborativa** como de [Tong et al.], donde los participantes buscan resolver dichos retos.

En el trabajo de [Lee et al.] el **análisis de datos urbanos se basa en al menos tres tareas directas**, complementadas con actividades de exploración, buscando fomentar estrategias de búsqueda colaborativa entre sus usuarios. Así pues, **mis hipótesis serán evaluadas para ver su complejidad al resolverlas** y también para ver si estas direccionarían a los usuarios a trabajar de manera colaborativa.



HIPÓTESIS

01

¿En qué vecindario (edificio/apartamento) vive la mayor cantidad de familias con niños y qué características hacen que esa área sea tan popular para ellos?

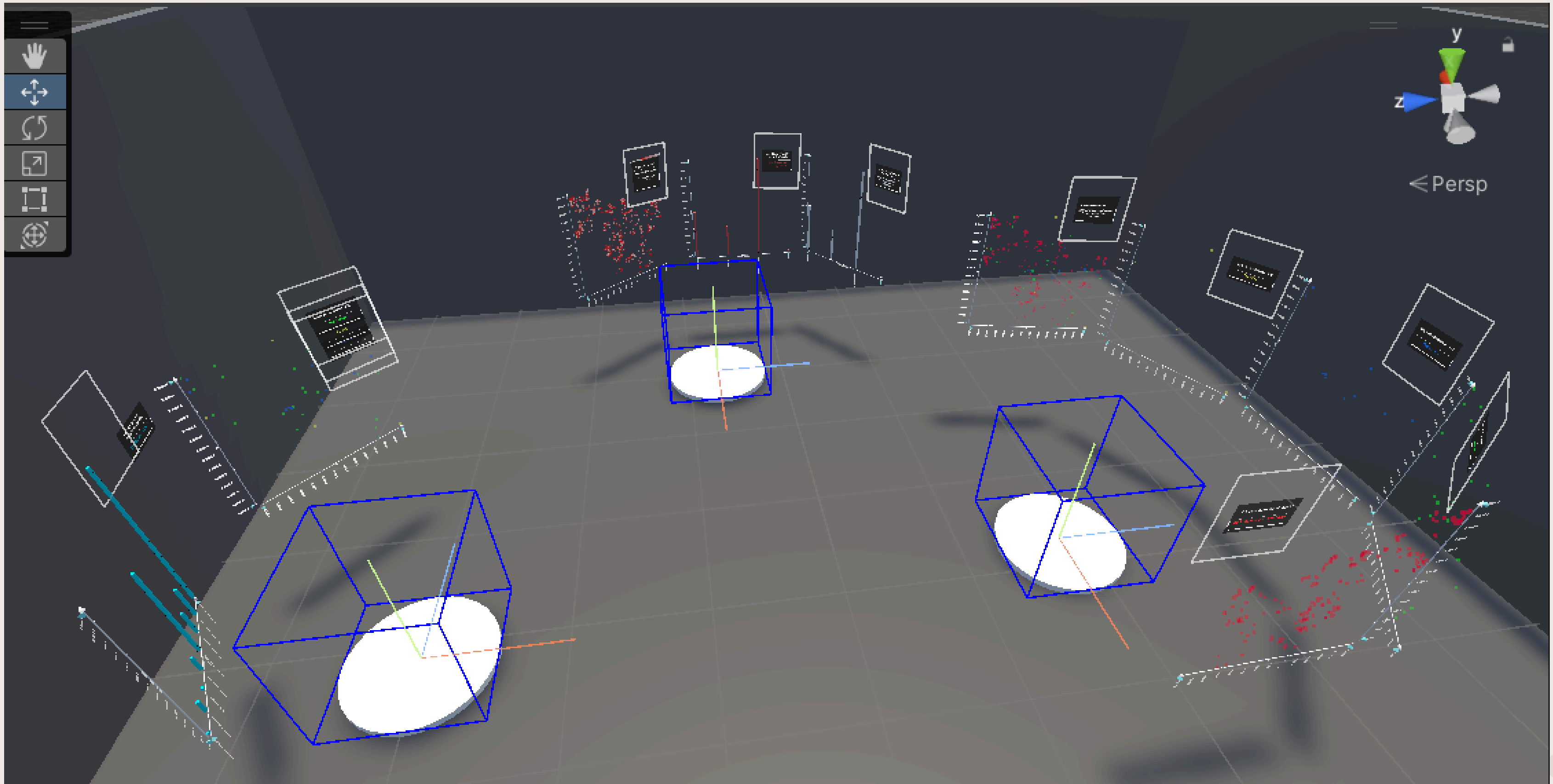
02

¿El nivel económico está relacionado con el grado de estudios y la ubicación geográfica?

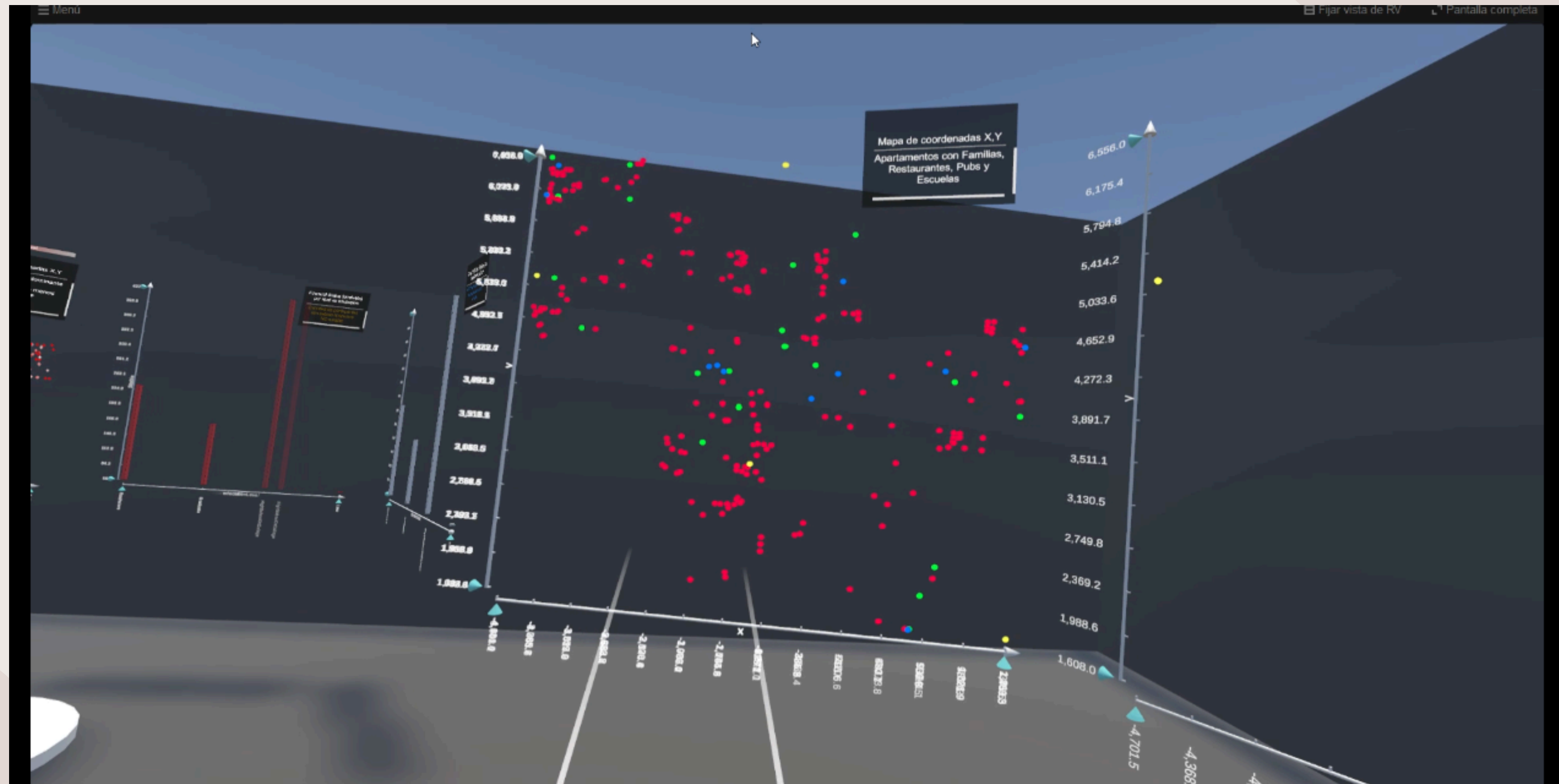
03

Imagina que una familia de cuatro personas dispone de poco más de 3000 USD al mes y al menos un miembro con educación secundaria. ¿En qué edificio o conjunto de apartamentos les recomendarías vivir y por qué esa opción es la mejor?

DASHBOARD:

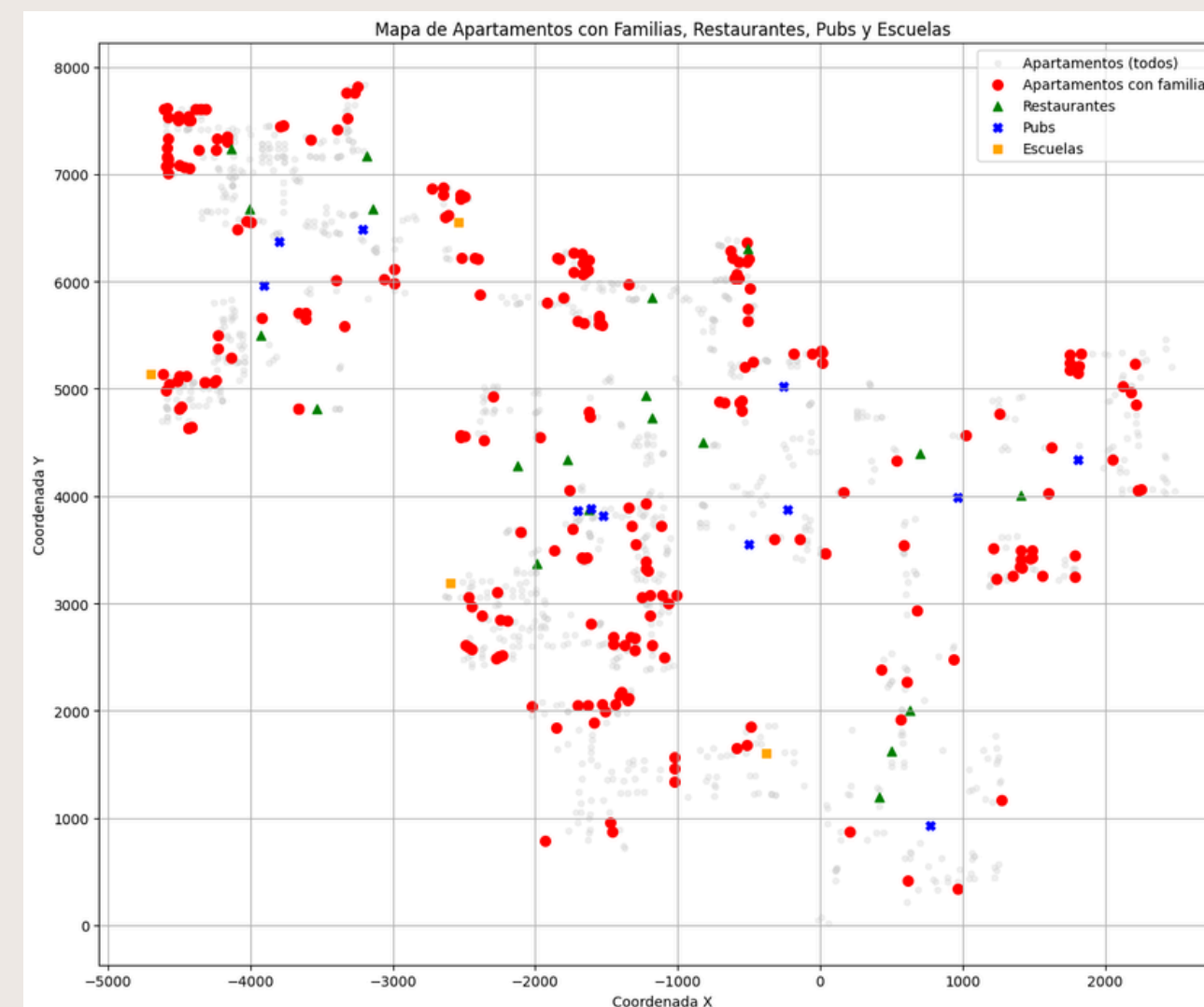


Hipótesis 1: ¿En qué vecindario (edificio/apartamento) vive la mayor cantidad de familias con niños y qué características hacen que esa área sea tan popular para ellos?

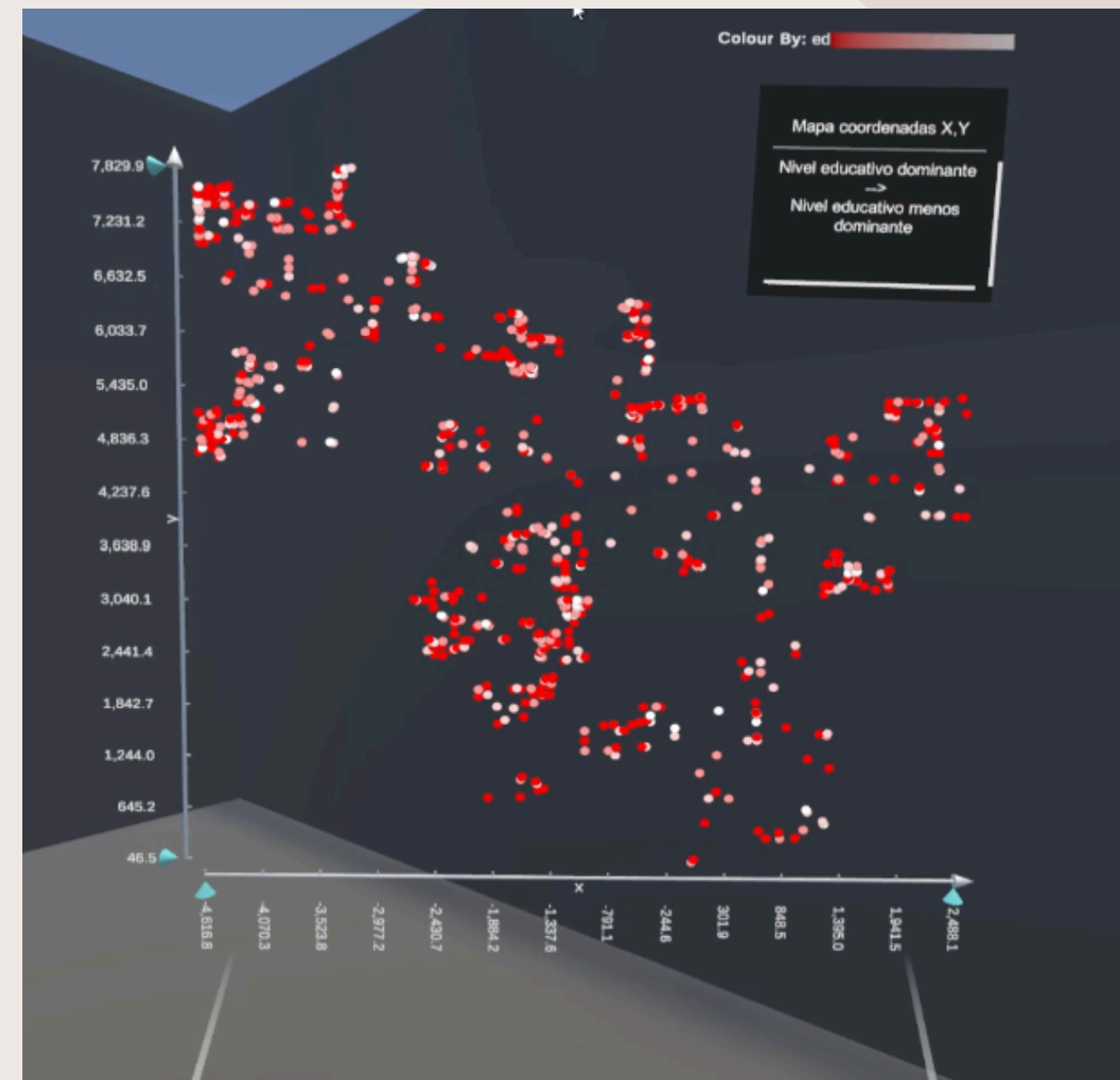
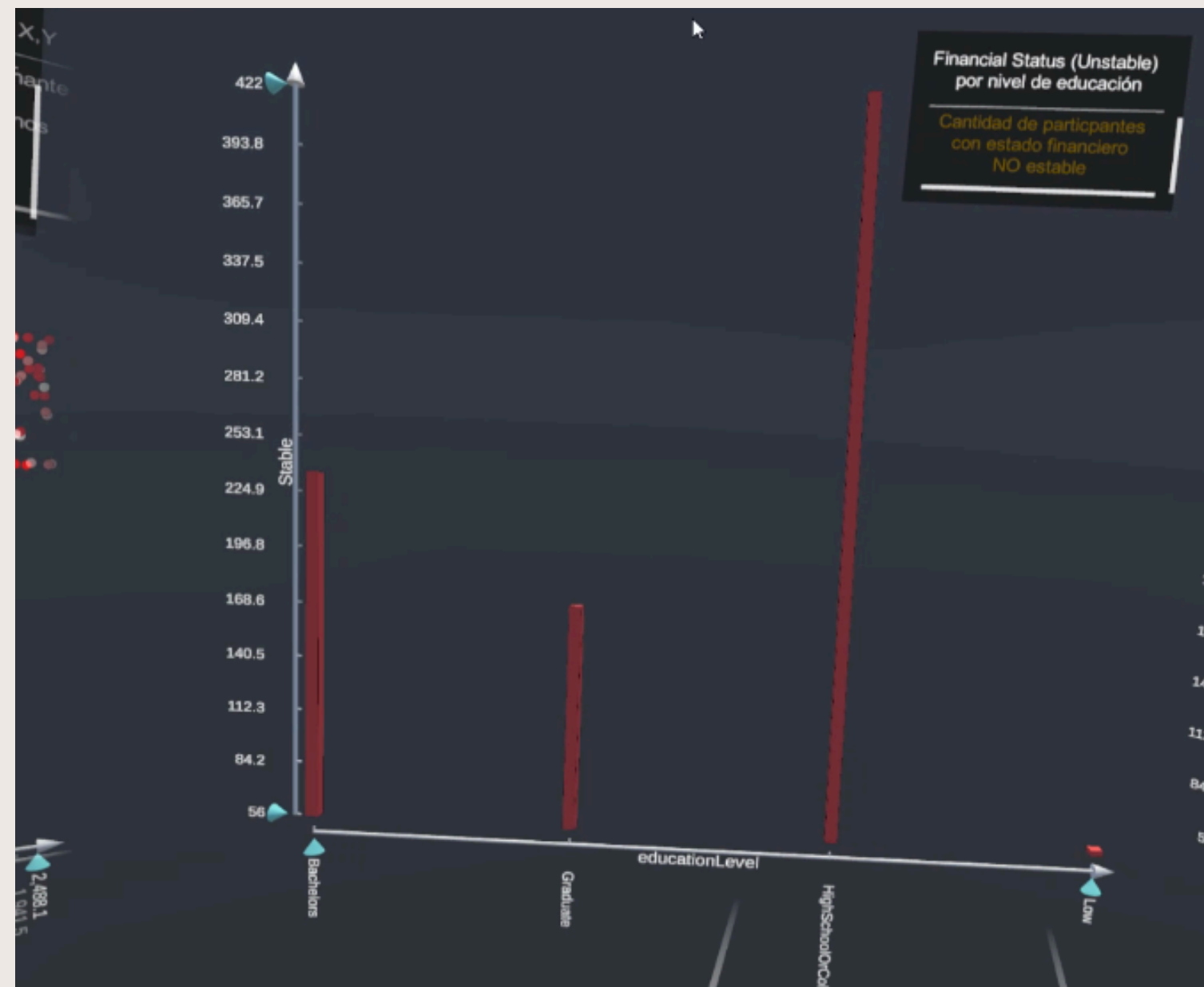


Hipótesis 1: ¿En qué vecindario (edificio/apartamento) vive la mayor cantidad de familias con niños y qué características hacen que esa área sea tan popular para ellos?

LA MAYOR CANTIDAD DE FAMILIAS CON NIÑOS VIVEN EN LOS EDIFICIOS QUE SE UBICAN EN LA PARTE SUPERIOR IZQUIERDA Y CASI EN LA PARTE CÉNTRICA DE LA CIUDAD. LAS CARACTERÍSTICAS DEL POR QUÉ ESTOS SE AGRUPAN MÁS EN ESTOS LUGARES ES PORQUE SE TIENE MAYOR PRESENCIA TANTO DE ESCUELAS (PRINCIPALMENTE PORQUE TIENEN NIÑOS), RESTAURANTES Y PUBS.



Hipótesis 2: ¿El nivel económico está relacionado con el grado de estudios y la ubicación geográfica?

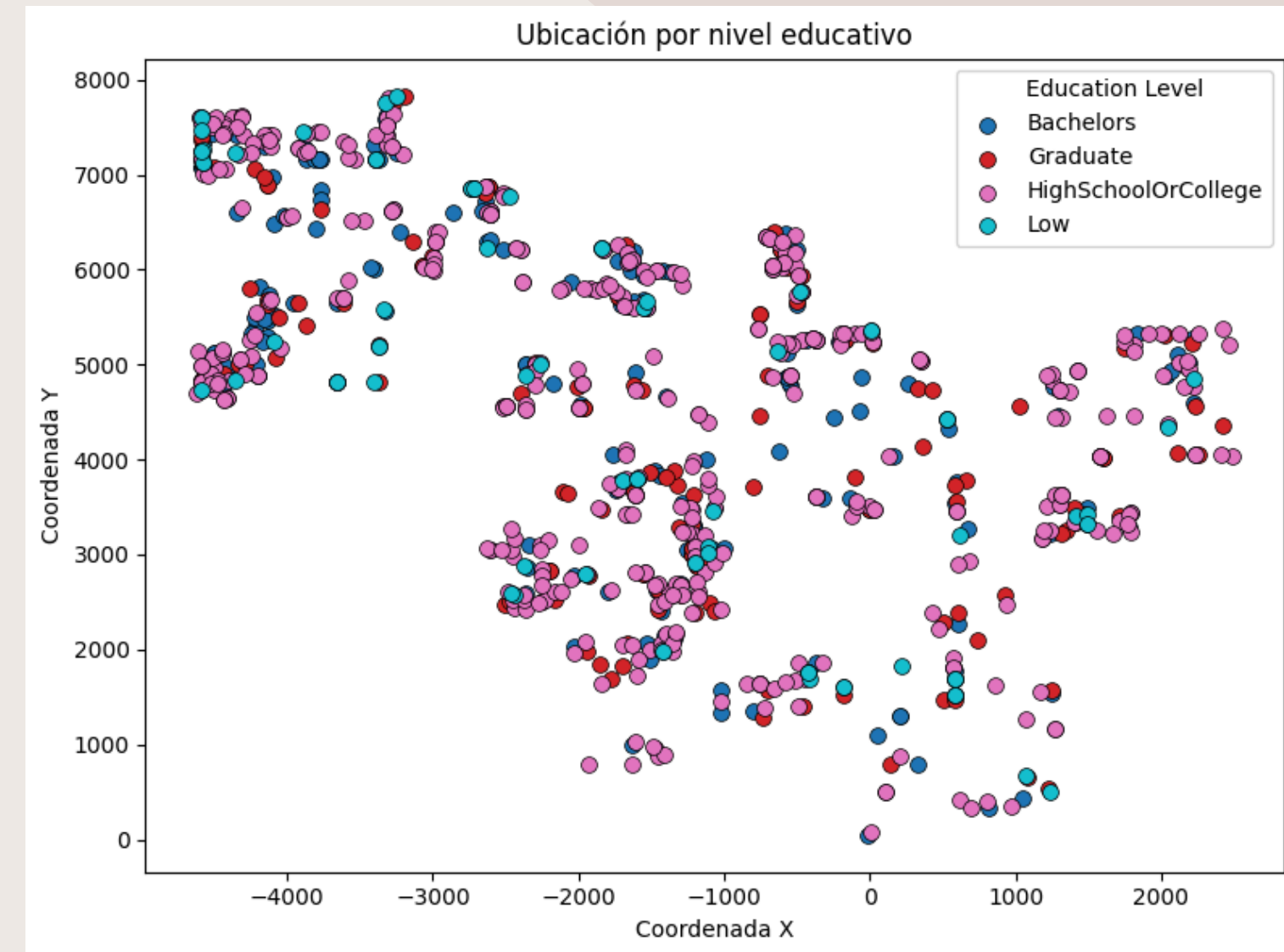


Hipótesis 2: ¿El nivel económico está relacionado con el grado de estudios y la ubicación geográfica?

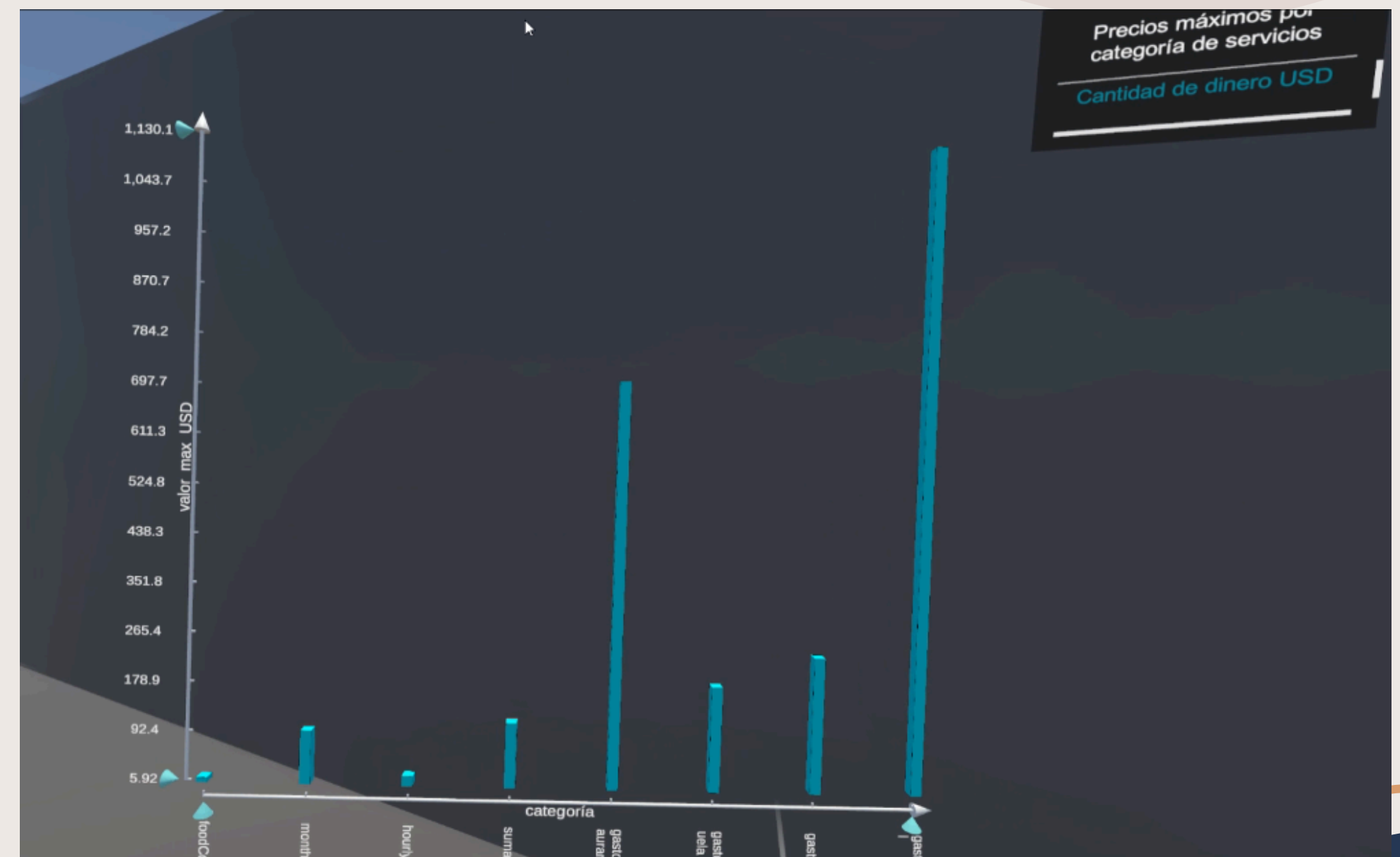
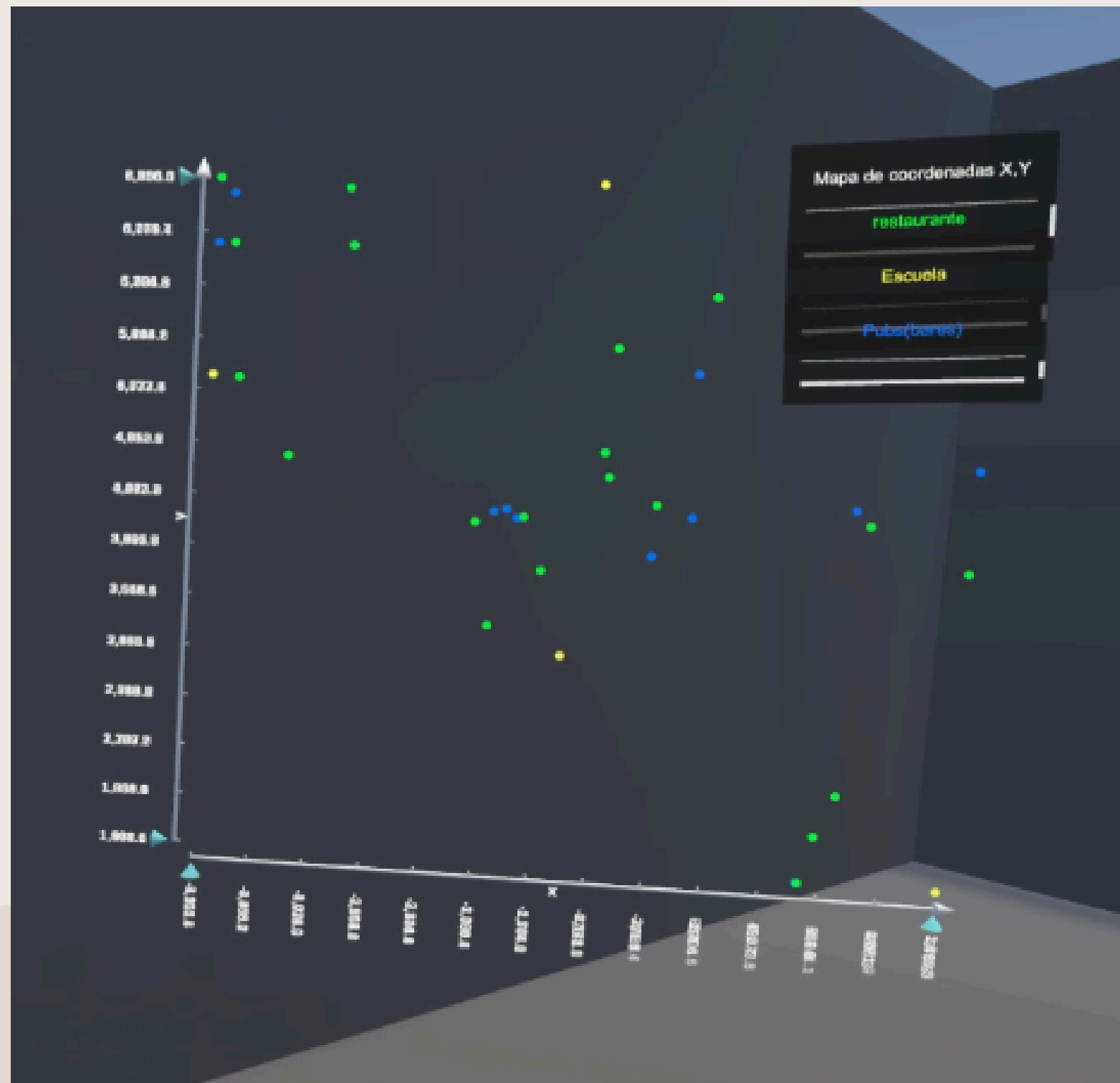
EN EL GRÁFICO DE BARRAS DE LA DISTRIBUCIÓN DE FINANCIALSTATUS POR NIVEL DE EDUCACIÓN, AHÍ TENEMOS QUE EL ESTADO FINANCIERO DE **"ESTABLE"** ES **LIDERADO POR HIGHSCHOOLORCOLLAGE CON MÁS DE 4000 PARTICIPANTES, PERO TAMBIÉN LIDERA EN "INESTABLE"**.

DANDO A CONOCER QUE SE PODRÍA DAR CAMBIOS ENTRE ESTOS 2 ESTADOS PARA DICHOS PARTICIPANTES. **QUIENES NO PRESENTAN EL ESTADO DE "INESTABLE", SON LOS QUE TIENEN GRADOS DE BACHILLER O SON GRADUADOS.**

MIENTRAS QUE SI OBSEVAMOS SU UBICACIÓN EN EL **MAPA**, PODEMOS APRECIAR QUE **LOS DE NIVEL HIGHSCHOOLORCOLLAGE, SE ENCUENTRAN ESPARCIDOS POR CASI TODA LA CIUDAD, MIENTRAS QUE SE TIENE MÁS PRESENCIA POR EL CENTRO DE LA CIUDAD A LOS PARTICIPANTES CON NIVEL DE ESTUDIO BACHILLER Y GRADUADO.** SI RECORDAMOS EL ESTADO FINANCIERO DE LOS PARTICIPANTES, **LOS DE NIVEL DE ESTUDIO HIGHSCHOOLORCOLLAGE, SON QUIENES TIENEN UN NIVEL DE "ESTABLE", LO CUAL PUEDE QUE MUEVA DIFERENTES NEGOCIOS EN LAS ZONAS DONDE MAYORMENTE HABITAN, JUNTO CON LOS DEMÁS PARTICIPANTES QUE TIENEN NIVEL FINANCIERO ESTABLE.** COMO EN EL **CENTRO DE LA CIUDAD Y EN LA ESQUINA SUPERIOR IZQUIERDA DEL MAPA.**



Hipótesis 3: Imagina que una familia de cuatro personas dispone de poco más de 3 000 USD al mes. ¿En qué edificio o conjunto de apartamentos les recomendarías vivir y por qué esa opción es la mejor?

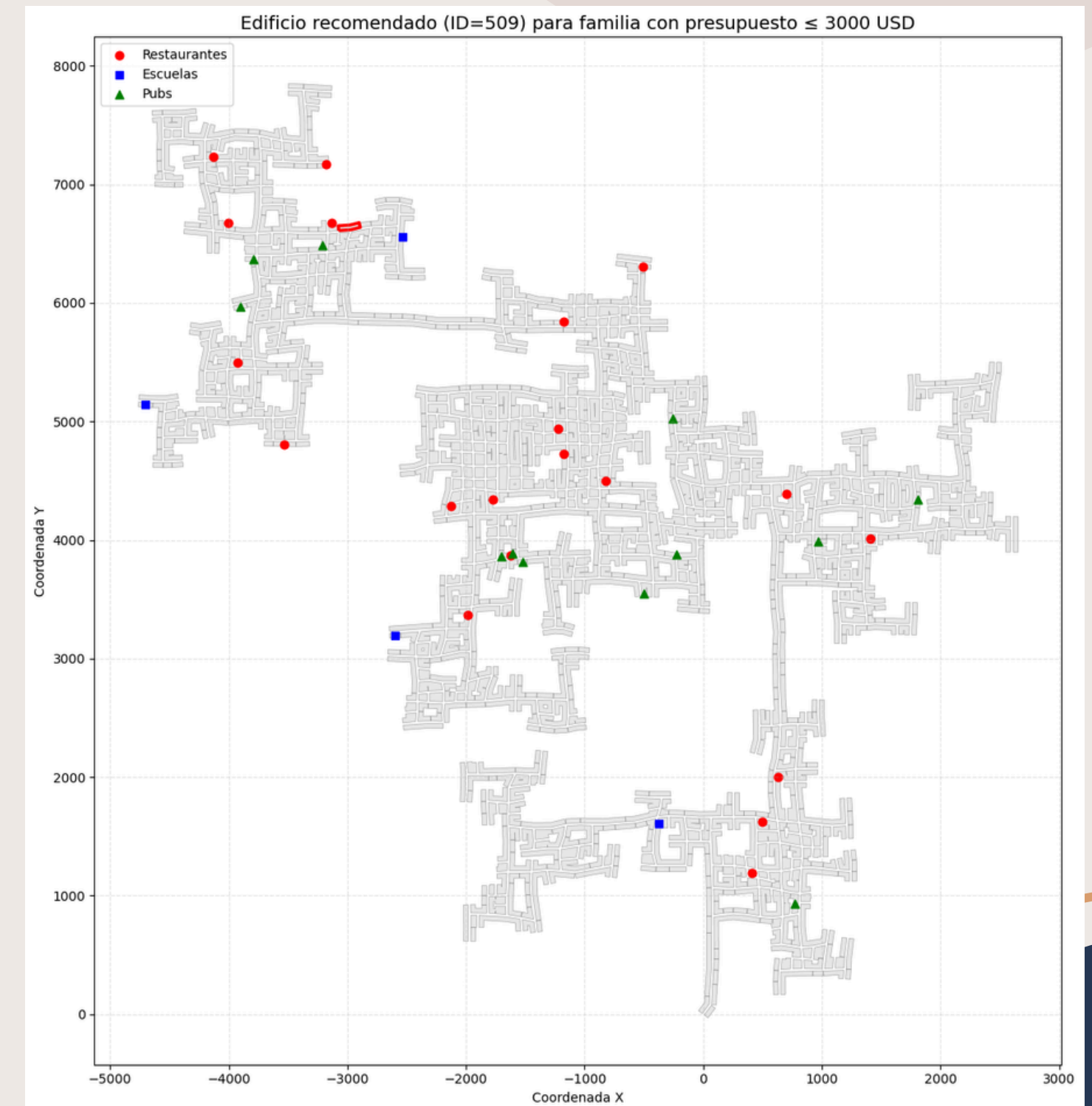


Hipótesis 3: Imagina que una familia de cuatro personas dispone de poco más de 3 000 USD al mes. ¿En qué edificio o conjunto de apartamentos les recomendarías vivir y por qué esa opción es la mejor?

DEBIDO A QUE LA FAMILIA TIENEN UN POCO MÁS DE 3000 DÓLARES, SI VEMOS LOS CÁLCULOS REALIZADOS EN EL CASO ANTERIOR:

CADA PERSONA COME 1 VEZ AL DÍA, OSEA 4×30 COMIDAS AL MES. 2 NIÑOS ASISTEN A LA ESCUELA Y PAGAN MENSUALIDAD. 2 ADULTOS VAN AL PUB 8 VECES AL MES, 1 HORA CADA VEZ. CON PAGO MÁXIMO. GASTO MENSUAL ESTIMADO: RESTAURANTE: 710.40 USD, ESCUELA: 182.29 USD, PUB: 237.45 USD, TOTAL: 1130.13 USD

PODEMOS CONCLUIR, QUE LA FAMILIA NO GASTARÍA MÁS DE SU PRESUPUESTO AL MES, INCLUSO SI PAGARAN LOS PRECIOS MÁS ELEVADOS. ADEMÁS, CONSIDERANDO QUE ES POSIBLE QUE LA FAMILIA TENGA NIÑOS, ES PREFERIBLE QUE VIVAN CERCA DE UNA ESCUELA. SI OBSERVAMOS EL GRÁFICO, SE TIENEN 4 ESCUELAS EN TOTAL, PERO TAMBIÉN ES PRIMORDIAL, EN SEGUNDO LUGAR, ESTAR CERCA DE RESTAURANTES. DONDE SI NOS DIRIGIMOS HACIA LA PARTE SUPERIOR IZQUIERDA DEL MAPA, TENEMOS LOS EDIFICIOS MÁS ADECUADOS PARA QUE ELLOS PUEDAN VIVIR SIN NINGÚN PROBLEMA, TENIENDO AL ALCANCE 4 RESTAURANTES, 3 PUBS Y 2 ESCUELAS. ADEMÁS EN EL ÚLTIMO GRÁFICO UBICAMOS EL MEJOR EDIFICIO PARA ELLOS TOMANDO EN CUENTA LA DISTANCIA MÍNIMA HACIA RESTAURANTE, ESCUELA Y BARES.



CONCLUSIONES

Al terminar este análisis exploratorio de datos, siento que no solo aprendí a manipular datasets algo complejos, sino también a pensar como un verdadero científico de datos, a la vez me di cuenta que no es del todo fácil.. Me di cuenta de que detrás de cada hipótesis hay una historia que se puede revelar con paciencia, lógica y herramientas adecuadas. El trabajo con los datos del VAST Challenge 2022 me permitió integrar conocimientos técnicos con el pensamiento crítico, y me mostró cómo la visualización y el análisis pueden ayudar a responder preguntas sociales importantes, aunque se trate de una ciudad ficticia.

También entendí que en el análisis de datos no todo es lineal: muchas veces tuve que volver sobre mis pasos, reevaluar decisiones y ajustar enfoques. Esto, lejos de ser una falla, es parte natural del proceso. Cabe resaltar que considero que aún continuo en una etapa de aprendizaje, ya que el campo de la ciencia de datos es muy extenso.



**MUCHAS
GRACIAS**

The background of the image is split into two main color areas. The left side is a solid dark blue, while the right side is a light beige. A wavy, hand-drawn style orange line separates the two colors, starting from the top right and curving down towards the bottom right. The text 'MUCHAS GRACIAS' is written in a bold, white, sans-serif font on the dark blue background.