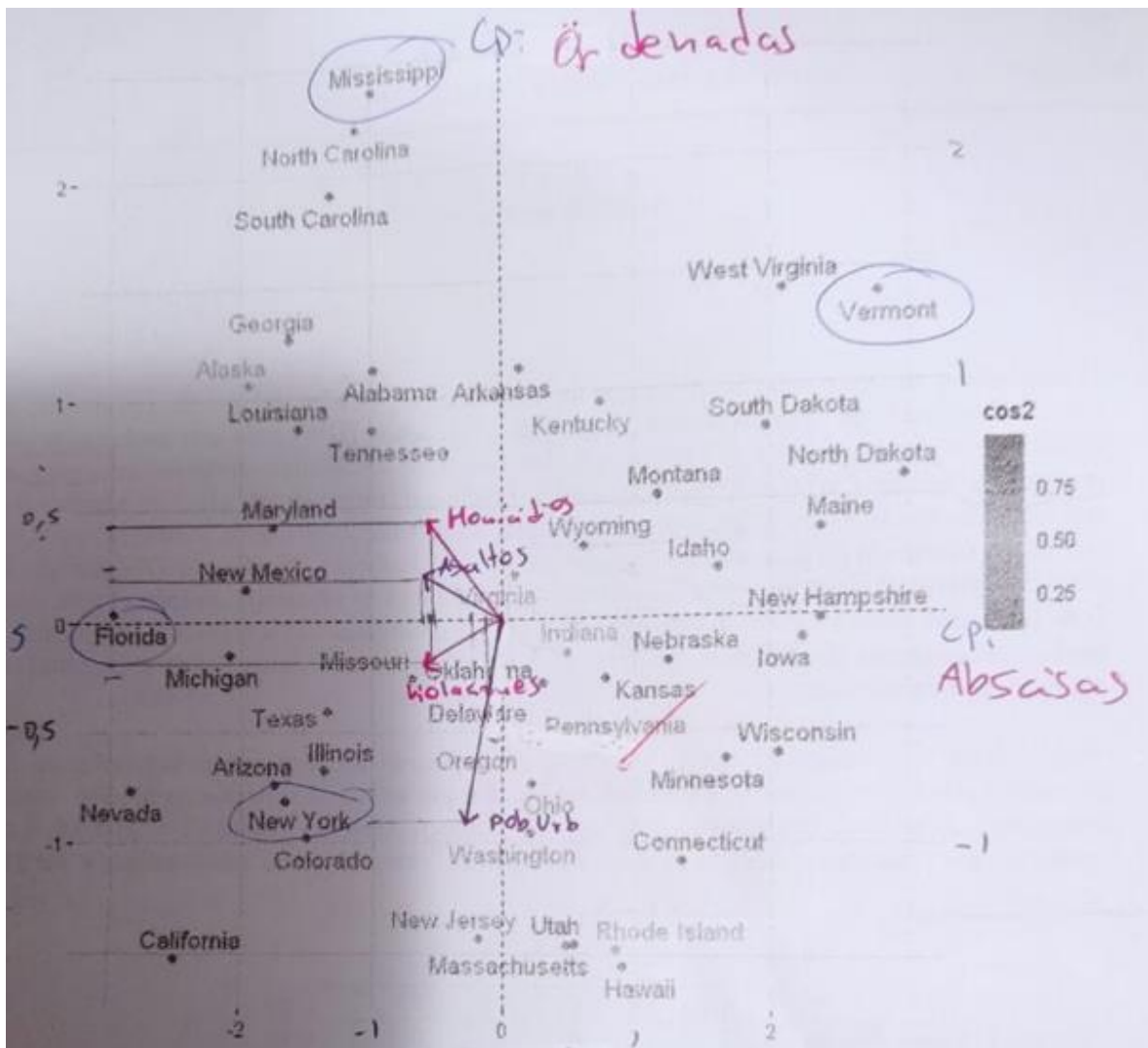


1. El mercado de automóviles chileno es uno de los más diversificados del mundo, con más de 60 marcas diferentes. Se requiere construir un Data Warehouse (DW) para una automotora que relacione los tipos de autos y sus clientes, los cuales se encuentran divididos en clases de preferencias como CarLovers, PadredeFamilia, Funcional, Empoderada, etc. Los automóviles por su parte, se clasifican en marcas agrupadas en: Europeos, Americanos, Asiáticos, etc. Que por su vez contienen características como: cedan, citycar, hushback, etc. En este DW también se debe considerar el lugar de venta: puesto que la automotora se encuentra en todo el país. Este DW permitirá analizar: costos, ventas y descuentos. Identifique dimensiones y las medidas, realice un esquema o esquemas gráficos de hipercubos en tres dimensiones y presente un ejemplo de dato atómico. (1.5)

2. Para realizar un estudio del crimen en USA, se analizaron las detenciones policiales en los diferentes estados americanos. Para el análisis se consideraron solo 4 variables: población urbana (PobUrb) y número de tres tipos de crímenes Violaciones, Asaltos y Homicidios. El análisis de componentes principales muestra la siguiente tabla de resultados, para componentes y valores propios. (2.5)

	CP1	CP2	CP3	CP4
Valor Propio	2,48	0,90	0,388	0,18
PobUrb	-0,2781	-0,8728	-0,3780	0,1338
Violaciones	-0,5434	-0,1673	0,8177	0,0890
Asaltos	-0,5831	0,1879	-0,2681	-0,7434
Homicidios	-0,5358	0,41818	-0,3412	0,6492

- Determine el porcentaje de pérdida de la información, en un análisis de dos componentes.
- En la gráfica de las dos primeras componentes, sobre los estados, grafique las variables originales en el plano (primera componente como la abscisa y la segunda como ordenada).
- Interprete cada una de las dos primeras componentes, con un solo concepto cada una.
- Interprete el clima criminal en estados como Florida, New York, Vermont y Mississippi.



3. Realice un dendrograma usando el método de agrupamiento jerárquico, de las ciudades españolas según su distancia, dada en la siguiente tabla. Indique cuales ciudades formarían dos y tres grupos. (1.0)

	Barcelona	Madrid	San Sebastián	Sevilla	Valencia
Barcelona	0	364	474	542	364
Madrid		0	474	542	355
San Sebastián			0	542	474
Sevilla				0	542
Valencia					0

4. Realice una analogía entre regresión lineal y regresión logística, explicando los parámetros para evaluar cada de las características de estas regresiones. (1.0)