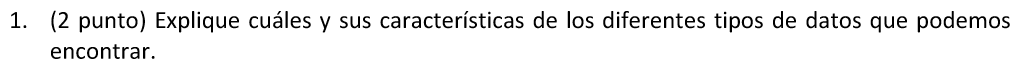


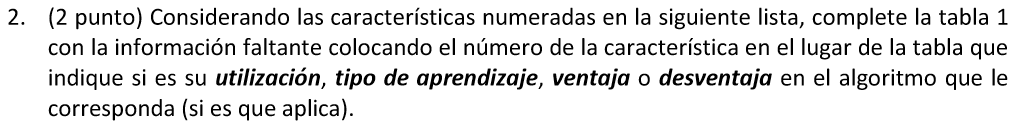
 Eduardo Castillo Martínez

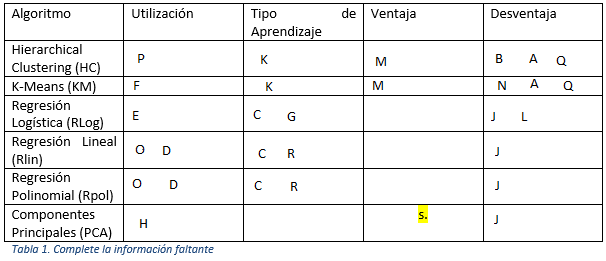


* Cuantitativos: datos numéricos que representan cierta jerarquía dependiendo su magnitud.
* Cualitativos:
* Multiestado: Representan clasificaciones de una variable.
* Enteros positivos: Normalmente son datos Multiestado transformados a

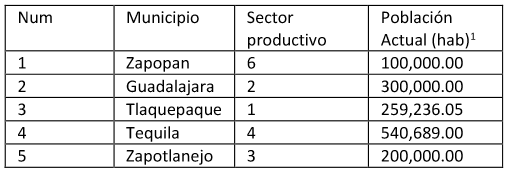
numéricos.

* Nominales: no representa jerarquía
* Ordinarios: dependiendo de la magnitud se establece jerarquía
* Binarios: datos con solo dos clasificaciones de tipo booleano, cierto o falso normalmente representando por 1 y 0.
* Genéticos: clasificación de especies.



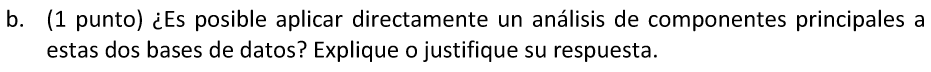


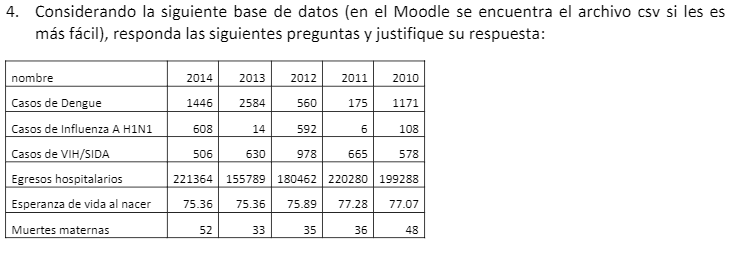


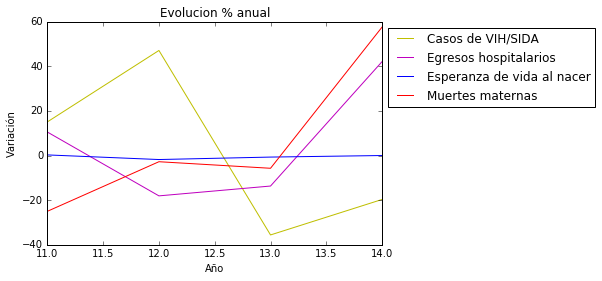


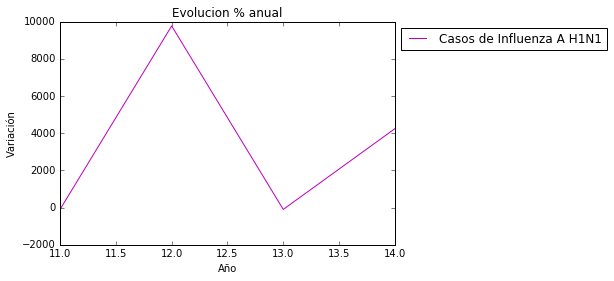


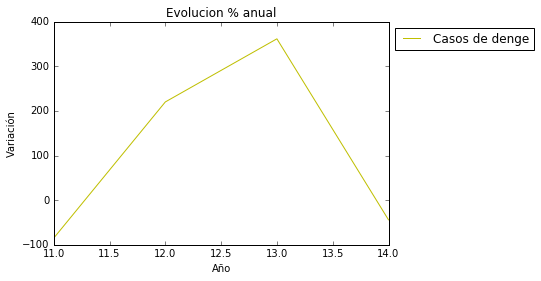
* Estado: Cualitativa, Multiestado
* Zona geográfica: Cualitativa, Enteros positivos nominales.
* Índice de delincuencia: Cuantitativa
* Municipio: Cualitativa, Multiestado
* Sector productivo: Cualitativa, Enteros positivos nominales.
* Población Actual: Cuantitativa

No es posible, el ACP requiere que las variables categóricas(Multiestado) sean convertidas en variables tipo “dummy”.







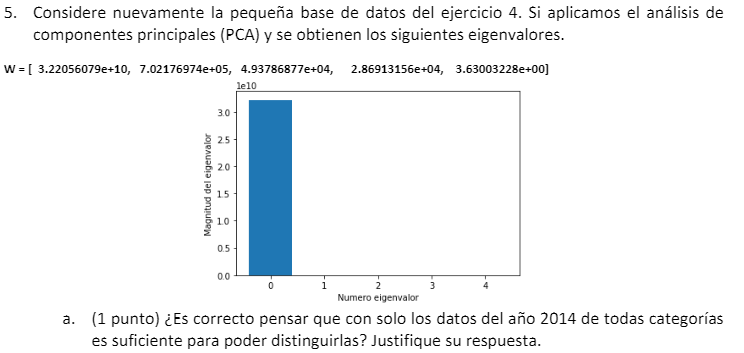




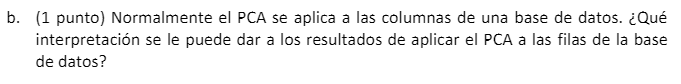
Casos de Influenza A H1N1



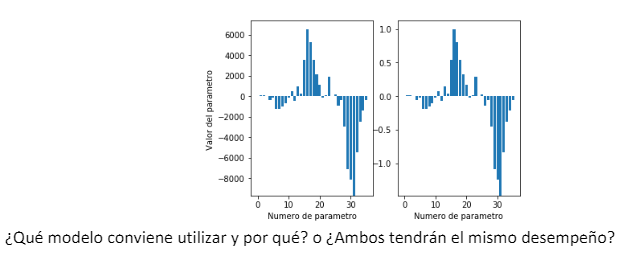
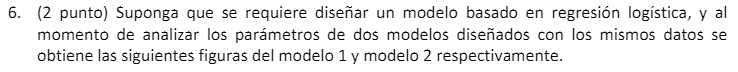
Basado en análisis visual de los gráficos, media y correlación, se determinó que las categorías con desempeño similar son" Muertes maternas" y "Egresos hospitalarios".



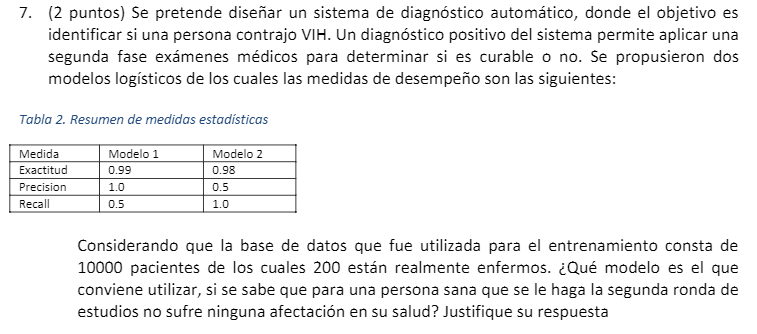
Si puede ser suficiente para distinguirla, dado que ese año es el eje donde se tiene una mayor dispersión de los datos



Al aplicar el ACP a las filas de la base de datos, se estaría buscando cuales son las categorías en donde se presenta la mayor dispersión de los datos, a diferencia del análisis de columnas dónde se busca cuales son los años con mayor dispersión, para el análisis de columnas se proyectan los años con menos dispersión en los años con mayor, en el caso de análisis de filas se proyectan las categorías de menor sobre las de mayor dispersión con el fin de reducir la dimensionalidad de los datos.



Los dos modelos tendrían el mismo desempeño, parece ser que el modelo 2 está estandarizado, pero aún así los coeficientes están asignados igual en los dos modelos proporcionalmente, los parámetros de mayor orden siguen teniendo una ponderación mayor en el modelo. Podría ser conveniente aplicar una regresión regularizada para evitar ese problema.

La exactitud no es un buen criterio en este caso, dado que el número de enfermos es pequeño en relación al número total de pacientes. Así el modelo clasifique que todos están sanos, la exactitud seguiría siendo alta.

En este caso es mejor enfocarse en el recall, si la precisión no es tan buena, la consecuencia sería que el modelo clasificaría como enfermo a pacientes que no lo están.

Al tener un mal desempeño en el recall no se estarían detectando pacientes que si están enfermos, causando posiblemente su muerte, por lo tanto es mejor utilizar para este caso el modelo 2.