|  |  |
| --- | --- |
|  | **UNIVERSIDAD DE LOS ANDES**  **DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**  **Maestría en inteligencia analítica para la toma de decisiones**  **Minería de Datos**  **Profesor Carlos Valencia** |

Nombre: Juan Sebastián Garcés C. Código: 201628949

Fecha: Junio 26 de 2018

## Quiz 1

Punto 1. Componentes Principales

Utilizando el conjunto de datos "airquality" del paquete "datasets", construya la matriz de datos X usando las primeras cuatro variables: nivel de ozono, radiación solar, viento y temperatura.

Información de los datos puede ser consultada ?airquality. (Use na.omit)

1. Suponga que se le pide construir un índice de severidad de clima para el mes de mayo usando estas cuatro variables. Encuentre la primera componente principal (con la matriz de varianza) e interprétela.

Se construye la matriz de covarianza teniendo en cuenta solo las 4 variables indicadas:

> # Matriz de covarianza

> S=round(var(air),2)

> S

Ozone Solar.R Wind Temp

Ozone 1107.29 1056.58 -72.51 221.52

Solar.R 1056.58 8308.74 -41.24 255.47

Wind -72.51 -41.24 12.66 -16.86

Temp 221.52 255.47 -16.86 90.82

dd