Predicting Chronic Kidney Disease

Michelle Koh

2024-10-06

Introduction

This is a multivariate dataset that I had found on UC Irvine's Machine Learning repository, which can be used to predict chronic kidney disease. Chronic kidney disease or CKD is usually identified with kidney damage or an estimated glomerular filtration rate of less than 60 mL/min/1.73 m^2 - usually consistently lower for 3 or more months.

This dataset contains 400 observations with 25 different variable, including: age, blood pressure, specific gravity, albumin, sugar, red blood cells, pus cell, pus cell clumps, bacteria, blood glucose random, blood urea, serum creatinine, sodium, potassium, hemoglobin, packed cell volume, white blood cell count, red blood cell count, hypertension, diabetes mellitus, coronary artery disease, appetite, pedal edema, anemia, class. All were collected between a 2 month period.

Based the results from our chronic kidney disease dataset, hopefully future chronic kidney disease patients will be identified, sooner the better.

Analysis Objectives

- To perform descriptive statistics and identify trends among the different variables and observations.
- Afterwards, visualize the trends by creating figures that is easier to follow.

Load the data

```
#Load the packages needed
library(pacman)
pacman::p_load(rio,
               dplyr, ggplot2, gtsummary)

#Import the dataset
kidney <- import("~/Desktop/data550/final_project/chronic_kidney_disease_data.csv") #to get this datase

#Displaying the first few rows of the dataset to get the f l a v o r of it
head(kidney)
```

```
##   age bp    sg al su    rbc       pc        pcc         ba bgr bu  sc sod pot
## 1  48 80 1.020  1  0          normal notpresent notpresent 121 36 1.2  NA  NA
## 2   7 50 1.020  4  0          normal notpresent notpresent  NA 18 0.8  NA  NA
## 3  62 80 1.010  2  3 normal   normal notpresent notpresent 423 53 1.8  NA  NA
## 4  48 70 1.005  4  0 normal abnormal    present notpresent 117 56 3.8 111 2.5
```

```
## 5  51 80 1.010  2  0 normal   normal notpresent notpresent 106 26 1.4  NA  NA
## 6  60 90 1.015  3  0                notpresent notpresent  74 25 1.1 142 3.2
##   hemo pcv wbcc rbcc htn  dm cad appet  pe ane target
## 1 15.4  44 7800  5.2 yes yes  no  good  no  no    ckd
## 2 11.3  38 6000   NA  no  no  no  good  no  no    ckd
## 3  9.6  31 7500   NA  no yes  no  poor  no yes    ckd
## 4 11.2  32 6700  3.9 yes  no  no  poor yes yes    ckd
## 5 11.6  35 7300  4.6  no  no  no  good  no  no    ckd
## 6 12.2  39 7800  4.4 yes yes  no  good yes  no    ckd
```

Clean dataset

```
#Let's see how many missing values there are
missing_summary <- kidney %>%
  summarise(across(everything(), ~ sum(is.na(.)) / n() * 100, .names = "missing_{col}"))

#Now print it
print(missing_summary)
```

```
##   missing_age missing_bp missing_sg missing_al missing_su missing_rbc
## 1        2.25          3      11.75       11.5      12.25           0
##   missing_pc missing_pcc missing_ba missing_bgr missing_bu missing_sc
## 1          0           0          0          11       4.75       4.25
##   missing_sod missing_pot missing_hemo missing_pcv missing_wbcc missing_rbcc
## 1       21.75          22           13       17.75         26.5        32.75
##   missing_htn missing_dm missing_cad missing_appet missing_pe missing_ane
## 1           0          0           0             0          0           0
##   missing_target
## 1              0
```

Descriptive tables

```
# Create summary table with gtsummary
characteristics_summary_kidney <- kidney %>%
  filter(target != "ckd\t") %>% #removing this extra ckd observation that was weirdly counted as its ow
  select( rbc, pc, pcc, ba, htn, dm, cad, appet, pe, ane, target) %>%  #filtering all the character var
  rename(
   `Red Blood Cells` = rbc,
   `Pus Cell` = pc,
   `Pus Cell Clumps` = pcc,
   Bacteria = ba,
   Hypertension = htn,
   Diabetes = dm,
   `Coronary Artery Disease` = cad,
   Appetite = appet,
   `Pedal Edema` = pe,
   Anemia = ane,
   CKD = target
  ) %>%
  tbl_summary(
```

```
    by = CKD,
    missing = "no"
  )

#Print the summary table
print(characteristics_summary_kidney)
```

```
## <div id="uzxdipybxd" style="padding-left:0px;padding-right:0px;padding-top:10px;padding-bottom:10px;
##   <style>#uzxdipybxd table {
##   font-family: system-ui, 'Segoe UI', Roboto, Helvetica, Arial, sans-serif, 'Apple Color Emoji', 'Se
##   -webkit-font-smoothing: antialiased;
##   -moz-osx-font-smoothing: grayscale;
## }
##
## #uzxdipybxd thead, #uzxdipybxd tbody, #uzxdipybxd tfoot, #uzxdipybxd tr, #uzxdipybxd td, #uzxdipybxd
##   border-style: none;
## }
##
## #uzxdipybxd p {
##   margin: 0;
##   padding: 0;
## }
##
## #uzxdipybxd .gt_table {
##   display: table;
##   border-collapse: collapse;
##   line-height: normal;
##   margin-left: auto;
##   margin-right: auto;
##   color: #333333;
##   font-size: 16px;
##   font-weight: normal;
##   font-style: normal;
##   background-color: #FFFFFF;
##   width: auto;
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #A8A8A8;
##   border-right-style: none;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #A8A8A8;
##   border-left-style: none;
##   border-left-width: 2px;
##   border-left-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_caption {
##   padding-top: 4px;
##   padding-bottom: 4px;
## }
```

```
##
## #uzxdipybxd .gt_title {
##   color: #333333;
##   font-size: 125%;
##   font-weight: initial;
##   padding-top: 4px;
##   padding-bottom: 4px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-bottom-color: #FFFFFF;
##   border-bottom-width: 0;
## }
##
## #uzxdipybxd .gt_subtitle {
##   color: #333333;
##   font-size: 85%;
##   font-weight: initial;
##   padding-top: 3px;
##   padding-bottom: 5px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-top-color: #FFFFFF;
##   border-top-width: 0;
## }
##
## #uzxdipybxd .gt_heading {
##   background-color: #FFFFFF;
##   text-align: center;
##   border-bottom-color: #FFFFFF;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_bottom_border {
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_col_headings {
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
```

```
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_col_heading {
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: normal;
##   text-transform: inherit;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
##   vertical-align: bottom;
##   padding-top: 5px;
##   padding-bottom: 6px;
##   padding-left: 5px;
##   padding-right: 5px;
##   overflow-x: hidden;
## }
##
## #uzxdipybxd .gt_column_spanner_outer {
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: normal;
##   text-transform: inherit;
##   padding-top: 0;
##   padding-bottom: 0;
##   padding-left: 4px;
##   padding-right: 4px;
## }
##
## #uzxdipybxd .gt_column_spanner_outer:first-child {
##   padding-left: 0;
## }
##
## #uzxdipybxd .gt_column_spanner_outer:last-child {
##   padding-right: 0;
## }
##
## #uzxdipybxd .gt_column_spanner {
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   vertical-align: bottom;
##   padding-top: 5px;
##   padding-bottom: 5px;
##   overflow-x: hidden;
##   display: inline-block;
##   width: 100%;
```

```
## }
##
## #uzxdipybxd .gt_spanner_row {
##   border-bottom-style: hidden;
## }
##
## #uzxdipybxd .gt_group_heading {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: initial;
##   text-transform: inherit;
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
##   vertical-align: middle;
##   text-align: left;
## }
##
## #uzxdipybxd .gt_empty_group_heading {
##   padding: 0.5px;
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: initial;
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   vertical-align: middle;
## }
##
## #uzxdipybxd .gt_from_md > :first-child {
##   margin-top: 0;
## }
##
## #uzxdipybxd .gt_from_md > :last-child {
##   margin-bottom: 0;
## }
```

```
##
## #uzxdipybxd .gt_row {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   margin: 10px;
##   border-top-style: solid;
##   border-top-width: 1px;
##   border-top-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
##   vertical-align: middle;
##   overflow-x: hidden;
## }
##
## #uzxdipybxd .gt_stub {
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: initial;
##   text-transform: inherit;
##   border-right-style: solid;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #uzxdipybxd .gt_stub_row_group {
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: initial;
##   text-transform: inherit;
##   border-right-style: solid;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
##   padding-left: 5px;
##   padding-right: 5px;
##   vertical-align: top;
## }
##
## #uzxdipybxd .gt_row_group_first td {
##   border-top-width: 2px;
## }
##
## #uzxdipybxd .gt_row_group_first th {
##   border-top-width: 2px;
## }
```

```
##
## #uzxdipybxd .gt_summary_row {
##   color: #333333;
##   background-color: #FFFFFF;
##   text-transform: inherit;
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #uzxdipybxd .gt_first_summary_row {
##   border-top-style: solid;
##   border-top-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_first_summary_row.thick {
##   border-top-width: 2px;
## }
##
## #uzxdipybxd .gt_last_summary_row {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_grand_summary_row {
##   color: #333333;
##   background-color: #FFFFFF;
##   text-transform: inherit;
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #uzxdipybxd .gt_first_grand_summary_row {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-top-style: double;
##   border-top-width: 6px;
##   border-top-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_last_grand_summary_row_top {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
```

```
##   padding-right: 5px;
##   border-bottom-style: double;
##   border-bottom-width: 6px;
##   border-bottom-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_striped {
##   background-color: rgba(128, 128, 128, 0.05);
## }
##
## #uzxdipybxd .gt_table_body {
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_footnotes {
##   color: #333333;
##   background-color: #FFFFFF;
##   border-bottom-style: none;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 2px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
## }
##
## #uzxdipybxd .gt_footnote {
##   margin: 0px;
##   font-size: 90%;
##   padding-top: 4px;
##   padding-bottom: 4px;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #uzxdipybxd .gt_sourcenotes {
##   color: #333333;
##   background-color: #FFFFFF;
##   border-bottom-style: none;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 2px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
```

```
## }
##
## #uzxdipybxd .gt_sourcenote {
##   font-size: 90%;
##   padding-top: 4px;
##   padding-bottom: 4px;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #uzxdipybxd .gt_left {
##   text-align: left;
## }
##
## #uzxdipybxd .gt_center {
##   text-align: center;
## }
##
## #uzxdipybxd .gt_right {
##   text-align: right;
##   font-variant-numeric: tabular-nums;
## }
##
## #uzxdipybxd .gt_font_normal {
##   font-weight: normal;
## }
##
## #uzxdipybxd .gt_font_bold {
##   font-weight: bold;
## }
##
## #uzxdipybxd .gt_font_italic {
##   font-style: italic;
## }
##
## #uzxdipybxd .gt_super {
##   font-size: 65%;
## }
##
## #uzxdipybxd .gt_footnote_marks {
##   font-size: 75%;
##   vertical-align: 0.4em;
##   position: initial;
## }
##
## #uzxdipybxd .gt_asterisk {
##   font-size: 100%;
##   vertical-align: 0;
## }
##
## #uzxdipybxd .gt_indent_1 {
##   text-indent: 5px;
## }
##
```

```
## #uzxdipybxd .gt_indent_2 {
##   text-indent: 10px;
## }
##
## #uzxdipybxd .gt_indent_3 {
##   text-indent: 15px;
## }
##
## #uzxdipybxd .gt_indent_4 {
##   text-indent: 20px;
## }
##
## #uzxdipybxd .gt_indent_5 {
##   text-indent: 25px;
## }
##
## #uzxdipybxd .katex-display {
##   display: inline-flex !important;
##   margin-bottom: 0.75em !important;
## }
##
## #uzxdipybxd div.Reactable > div.rt-table > div.rt-thead > div.rt-tr.rt-tr-group-header > div.rt-th-g
##   height: 0px !important;
## }
## </style>
##   <table class="gt_table" data-quarto-disable-processing="false" data-quarto-bootstrap="false">
##   <thead>
##     <tr class="gt_col_headings">
##       <th class="gt_col_heading gt_columns_bottom_border gt_left" rowspan="1" colspan="1" scope="col
##       <th class="gt_col_heading gt_columns_bottom_border gt_center" rowspan="1" colspan="1" scope="c
## N = 248</span><span class="gt_footnote_marks" style="white-space:nowrap;font-style:italic;font-weigh
##       <th class="gt_col_heading gt_columns_bottom_border gt_center" rowspan="1" colspan="1" scope="c
## N = 150</span><span class="gt_footnote_marks" style="white-space:nowrap;font-style:italic;font-weigh
##     </tr>
##   </thead>
##   <tbody class="gt_table_body">
##     <tr><td headers="label" class="gt_row gt_left">Red Blood Cells</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">142 (57%)</td>
## <td headers="stat_2" class="gt_row gt_center">9 (6.0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    abnormal</td>
## <td headers="stat_1" class="gt_row gt_center">47 (19%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    normal</td>
## <td headers="stat_1" class="gt_row gt_center">59 (24%)</td>
## <td headers="stat_2" class="gt_row gt_center">141 (94%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Pus Cell</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">55 (22%)</td>
## <td headers="stat_2" class="gt_row gt_center">9 (6.0%)</td></tr>
```

```
##     <tr><td headers="label" class="gt_row gt_left">    abnormal</td>
## <td headers="stat_1" class="gt_row gt_center">75 (30%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    normal</td>
## <td headers="stat_1" class="gt_row gt_center">118 (48%)</td>
## <td headers="stat_2" class="gt_row gt_center">141 (94%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Pus Cell Clumps</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">0 (0%)</td>
## <td headers="stat_2" class="gt_row gt_center">4 (2.7%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    notpresent</td>
## <td headers="stat_1" class="gt_row gt_center">207 (83%)</td>
## <td headers="stat_2" class="gt_row gt_center">146 (97%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    present</td>
## <td headers="stat_1" class="gt_row gt_center">41 (17%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Bacteria</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">0 (0%)</td>
## <td headers="stat_2" class="gt_row gt_center">4 (2.7%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    notpresent</td>
## <td headers="stat_1" class="gt_row gt_center">226 (91%)</td>
## <td headers="stat_2" class="gt_row gt_center">146 (97%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    present</td>
## <td headers="stat_1" class="gt_row gt_center">22 (8.9%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Hypertension</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">0 (0%)</td>
## <td headers="stat_2" class="gt_row gt_center">2 (1.3%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    no</td>
## <td headers="stat_1" class="gt_row gt_center">103 (42%)</td>
## <td headers="stat_2" class="gt_row gt_center">148 (99%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    yes</td>
## <td headers="stat_1" class="gt_row gt_center">145 (58%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Diabetes</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">0 (0%)</td>
## <td headers="stat_2" class="gt_row gt_center">2 (1.3%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">      no</td>
## <td headers="stat_1" class="gt_row gt_center">1 (0.4%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    no</td>
## <td headers="stat_1" class="gt_row gt_center">112 (45%)</td>
## <td headers="stat_2" class="gt_row gt_center">148 (99%)</td></tr>
```

```
##     <tr><td headers="label" class="gt_row gt_left">    yes</td>
## <td headers="stat_1" class="gt_row gt_center">135 (54%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Coronary Artery Disease</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">0 (0%)</td>
## <td headers="stat_2" class="gt_row gt_center">2 (1.3%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    no</td>
## <td headers="stat_1" class="gt_row gt_center">214 (86%)</td>
## <td headers="stat_2" class="gt_row gt_center">148 (99%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    yes</td>
## <td headers="stat_1" class="gt_row gt_center">34 (14%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Appetite</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">0 (0%)</td>
## <td headers="stat_2" class="gt_row gt_center">1 (0.7%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    good</td>
## <td headers="stat_1" class="gt_row gt_center">168 (68%)</td>
## <td headers="stat_2" class="gt_row gt_center">149 (99%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    poor</td>
## <td headers="stat_1" class="gt_row gt_center">80 (32%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Pedal Edema</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">0 (0%)</td>
## <td headers="stat_2" class="gt_row gt_center">1 (0.7%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    no</td>
## <td headers="stat_1" class="gt_row gt_center">172 (69%)</td>
## <td headers="stat_2" class="gt_row gt_center">149 (99%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    yes</td>
## <td headers="stat_1" class="gt_row gt_center">76 (31%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Anemia</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    </td>
## <td headers="stat_1" class="gt_row gt_center">0 (0%)</td>
## <td headers="stat_2" class="gt_row gt_center">1 (0.7%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    no</td>
## <td headers="stat_1" class="gt_row gt_center">189 (76%)</td>
## <td headers="stat_2" class="gt_row gt_center">149 (99%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    yes</td>
## <td headers="stat_1" class="gt_row gt_center">59 (24%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##   </tbody>
##
##   <tfoot class="gt_footnotes">
```

```
##     <tr>
##       <td class="gt_footnote" colspan="3"><span class="gt_footnote_marks" style="white-space:nowrap;i
##     </tr>
##   </tfoot>
## </table>
## </div>
```

```
# Numeric variables summary
numeric_summary<- kidney %>%
  filter(target != "ckd\t") %>%  #I removed this extra ckd observation that had only 2 observations
  select(age, bp, sg, bgr, bu, sc, sod, pot, hemo, pcv, wbcc, target) %>%
  rename(
    Age = age,
    `Blood Pressure (mm/Hg)` = bp,
    `Specific Gravity` = sg,
    `Blood Glucose Random (mgs/dl)` = bgr,
    `Blood Urea (mgs/dl)` = bu,
    `Serum Creatinine (mgs/dl)` =sc,
    `Sodium (mEq/L)` = sod,
    `Potassium (mEq/L)` = pot,
    `Hemoglobin (gms)` = hemo,
    `Packed Cell Volume` = pcv,
    `White Blood Cell Count` = wbcc,
    CKD = target
  ) %>%
  tbl_summary(
    by = CKD,
    missing = "no"
  )

# Print numeric summary table
print(numeric_summary)
```

```
## <div id="nmgucavudq" style="padding-left:0px;padding-right:0px;padding-top:10px;padding-bottom:10px;
##   <style>#nmgucavudq table {
##   font-family: system-ui, 'Segoe UI', Roboto, Helvetica, Arial, sans-serif, 'Apple Color Emoji', 'Se
##   -webkit-font-smoothing: antialiased;
##   -moz-osx-font-smoothing: grayscale;
## }
##
## #nmgucavudq thead, #nmgucavudq tbody, #nmgucavudq tfoot, #nmgucavudq tr, #nmgucavudq td, #nmgucavudq
##   border-style: none;
## }
##
## #nmgucavudq p {
##   margin: 0;
##   padding: 0;
## }
##
## #nmgucavudq .gt_table {
##   display: table;
##   border-collapse: collapse;
##   line-height: normal;
##   margin-left: auto;
```

```
##   margin-right: auto;
##   color: #333333;
##   font-size: 16px;
##   font-weight: normal;
##   font-style: normal;
##   background-color: #FFFFFF;
##   width: auto;
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #A8A8A8;
##   border-right-style: none;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #A8A8A8;
##   border-left-style: none;
##   border-left-width: 2px;
##   border-left-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_caption {
##   padding-top: 4px;
##   padding-bottom: 4px;
## }
##
## #nmgucavudq .gt_title {
##   color: #333333;
##   font-size: 125%;
##   font-weight: initial;
##   padding-top: 4px;
##   padding-bottom: 4px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-bottom-color: #FFFFFF;
##   border-bottom-width: 0;
## }
##
## #nmgucavudq .gt_subtitle {
##   color: #333333;
##   font-size: 85%;
##   font-weight: initial;
##   padding-top: 3px;
##   padding-bottom: 5px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-top-color: #FFFFFF;
##   border-top-width: 0;
## }
##
## #nmgucavudq .gt_heading {
##   background-color: #FFFFFF;
##   text-align: center;
##   border-bottom-color: #FFFFFF;
```

```
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_bottom_border {
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_col_headings {
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_col_heading {
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: normal;
##   text-transform: inherit;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
##   vertical-align: bottom;
##   padding-top: 5px;
##   padding-bottom: 6px;
##   padding-left: 5px;
##   padding-right: 5px;
##   overflow-x: hidden;
## }
##
## #nmgucavudq .gt_column_spanner_outer {
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: normal;
```

```
##   text-transform: inherit;
##   padding-top: 0;
##   padding-bottom: 0;
##   padding-left: 4px;
##   padding-right: 4px;
## }
##
## #nmgucavudq .gt_column_spanner_outer:first-child {
##   padding-left: 0;
## }
##
## #nmgucavudq .gt_column_spanner_outer:last-child {
##   padding-right: 0;
## }
##
## #nmgucavudq .gt_column_spanner {
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   vertical-align: bottom;
##   padding-top: 5px;
##   padding-bottom: 5px;
##   overflow-x: hidden;
##   display: inline-block;
##   width: 100%;
## }
##
## #nmgucavudq .gt_spanner_row {
##   border-bottom-style: hidden;
## }
##
## #nmgucavudq .gt_group_heading {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: initial;
##   text-transform: inherit;
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
##   vertical-align: middle;
```

```
##   text-align: left;
## }
##
## #nmgucavudq .gt_empty_group_heading {
##   padding: 0.5px;
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: initial;
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   vertical-align: middle;
## }
##
## #nmgucavudq .gt_from_md > :first-child {
##   margin-top: 0;
## }
##
## #nmgucavudq .gt_from_md > :last-child {
##   margin-bottom: 0;
## }
##
## #nmgucavudq .gt_row {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   margin: 10px;
##   border-top-style: solid;
##   border-top-width: 1px;
##   border-top-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 1px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 1px;
##   border-right-color: #D3D3D3;
##   vertical-align: middle;
##   overflow-x: hidden;
## }
##
## #nmgucavudq .gt_stub {
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: initial;
##   text-transform: inherit;
##   border-right-style: solid;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
```

```
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #nmgucavudq .gt_stub_row_group {
##   color: #333333;
##   background-color: #FFFFFF;
##   font-size: 100%;
##   font-weight: initial;
##   text-transform: inherit;
##   border-right-style: solid;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
##   padding-left: 5px;
##   padding-right: 5px;
##   vertical-align: top;
## }
##
## #nmgucavudq .gt_row_group_first td {
##   border-top-width: 2px;
## }
##
## #nmgucavudq .gt_row_group_first th {
##   border-top-width: 2px;
## }
##
## #nmgucavudq .gt_summary_row {
##   color: #333333;
##   background-color: #FFFFFF;
##   text-transform: inherit;
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #nmgucavudq .gt_first_summary_row {
##   border-top-style: solid;
##   border-top-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_first_summary_row.thick {
##   border-top-width: 2px;
## }
##
## #nmgucavudq .gt_last_summary_row {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
## }
```

```
##
## #nmgucavudq .gt_grand_summary_row {
##   color: #333333;
##   background-color: #FFFFFF;
##   text-transform: inherit;
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #nmgucavudq .gt_first_grand_summary_row {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-top-style: double;
##   border-top-width: 6px;
##   border-top-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_last_grand_summary_row_top {
##   padding-top: 8px;
##   padding-bottom: 8px;
##   padding-left: 5px;
##   padding-right: 5px;
##   border-bottom-style: double;
##   border-bottom-width: 6px;
##   border-bottom-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_striped {
##   background-color: rgba(128, 128, 128, 0.05);
## }
##
## #nmgucavudq .gt_table_body {
##   border-top-style: solid;
##   border-top-width: 2px;
##   border-top-color: #D3D3D3;
##   border-bottom-style: solid;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_footnotes {
##   color: #333333;
##   background-color: #FFFFFF;
##   border-bottom-style: none;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 2px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
```

```
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_footnote {
##   margin: 0px;
##   font-size: 90%;
##   padding-top: 4px;
##   padding-bottom: 4px;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #nmgucavudq .gt_sourcenotes {
##   color: #333333;
##   background-color: #FFFFFF;
##   border-bottom-style: none;
##   border-bottom-width: 2px;
##   border-bottom-color: #D3D3D3;
##   border-left-style: none;
##   border-left-width: 2px;
##   border-left-color: #D3D3D3;
##   border-right-style: none;
##   border-right-width: 2px;
##   border-right-color: #D3D3D3;
## }
##
## #nmgucavudq .gt_sourcenote {
##   font-size: 90%;
##   padding-top: 4px;
##   padding-bottom: 4px;
##   padding-left: 5px;
##   padding-right: 5px;
## }
##
## #nmgucavudq .gt_left {
##   text-align: left;
## }
##
## #nmgucavudq .gt_center {
##   text-align: center;
## }
##
## #nmgucavudq .gt_right {
##   text-align: right;
##   font-variant-numeric: tabular-nums;
## }
##
## #nmgucavudq .gt_font_normal {
##   font-weight: normal;
## }
##
## #nmgucavudq .gt_font_bold {
##   font-weight: bold;
```

```
## }
##
## #nmgucavudq .gt_font_italic {
##   font-style: italic;
## }
##
## #nmgucavudq .gt_super {
##   font-size: 65%;
## }
##
## #nmgucavudq .gt_footnote_marks {
##   font-size: 75%;
##   vertical-align: 0.4em;
##   position: initial;
## }
##
## #nmgucavudq .gt_asterisk {
##   font-size: 100%;
##   vertical-align: 0;
## }
##
## #nmgucavudq .gt_indent_1 {
##   text-indent: 5px;
## }
##
## #nmgucavudq .gt_indent_2 {
##   text-indent: 10px;
## }
##
## #nmgucavudq .gt_indent_3 {
##   text-indent: 15px;
## }
##
## #nmgucavudq .gt_indent_4 {
##   text-indent: 20px;
## }
##
## #nmgucavudq .gt_indent_5 {
##   text-indent: 25px;
## }
##
## #nmgucavudq .katex-display {
##   display: inline-flex !important;
##   margin-bottom: 0.75em !important;
## }
##
## #nmgucavudq div.Reactable > div.rt-table > div.rt-thead > div.rt-tr.rt-tr-group-header > div.rt-th-gr
##   height: 0px !important;
## }
## </style>
##   <table class="gt_table" data-quarto-disable-processing="false" data-quarto-bootstrap="false">
##   <thead>
##     <tr class="gt_col_headings">
##       <th class="gt_col_heading gt_columns_bottom_border gt_left" rowspan="1" colspan="1" scope="col
```

```
##       <th class="gt_col_heading gt_columns_bottom_border gt_center" rowspan="1" colspan="1" scope="c
## N = 248</span><span class="gt_footnote_marks" style="white-space:nowrap;font-style:italic;font-weigh
##       <th class="gt_col_heading gt_columns_bottom_border gt_center" rowspan="1" colspan="1" scope="c
## N = 150</span><span class="gt_footnote_marks" style="white-space:nowrap;font-style:italic;font-weigh
##     </tr>
##   </thead>
##   <tbody class="gt_table_body">
##     <tr><td headers="label" class="gt_row gt_left">Age</td>
## <td headers="stat_1" class="gt_row gt_center">59 (47, 66)</td>
## <td headers="stat_2" class="gt_row gt_center">46 (34, 58)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Blood Pressure (mm/Hg)</td>
## <td headers="stat_1" class="gt_row gt_center">80 (70, 90)</td>
## <td headers="stat_2" class="gt_row gt_center">70 (60, 80)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Specific Gravity</td>
## <td headers="stat_1" class="gt_row gt_center"><br /></td>
## <td headers="stat_2" class="gt_row gt_center"><br /></td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    1.005</td>
## <td headers="stat_1" class="gt_row gt_center">7 (3.4%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    1.01</td>
## <td headers="stat_1" class="gt_row gt_center">83 (40%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    1.015</td>
## <td headers="stat_1" class="gt_row gt_center">75 (36%)</td>
## <td headers="stat_2" class="gt_row gt_center">0 (0%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    1.02</td>
## <td headers="stat_1" class="gt_row gt_center">31 (15%)</td>
## <td headers="stat_2" class="gt_row gt_center">75 (52%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">    1.025</td>
## <td headers="stat_1" class="gt_row gt_center">11 (5.3%)</td>
## <td headers="stat_2" class="gt_row gt_center">70 (48%)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Blood Glucose Random (mgs/dl)</td>
## <td headers="stat_1" class="gt_row gt_center">144 (106, 220)</td>
## <td headers="stat_2" class="gt_row gt_center">108 (94, 124)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Blood Urea (mgs/dl)</td>
## <td headers="stat_1" class="gt_row gt_center">53 (32, 94)</td>
## <td headers="stat_2" class="gt_row gt_center">33 (24, 44)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Serum Creatinine (mgs/dl)</td>
## <td headers="stat_1" class="gt_row gt_center">2.3 (1.4, 4.6)</td>
## <td headers="stat_2" class="gt_row gt_center">0.9 (0.6, 1.1)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Sodium (mEq/L)</td>
## <td headers="stat_1" class="gt_row gt_center">136 (132, 139)</td>
## <td headers="stat_2" class="gt_row gt_center">141 (138, 146)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Potassium (mEq/L)</td>
## <td headers="stat_1" class="gt_row gt_center">4.30 (3.80, 4.90)</td>
## <td headers="stat_2" class="gt_row gt_center">4.50 (3.70, 4.90)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Hemoglobin (gms)</td>
## <td headers="stat_1" class="gt_row gt_center">10.90 (9.40, 12.10)</td>
## <td headers="stat_2" class="gt_row gt_center">15.00 (14.10, 16.20)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">Packed Cell Volume</td>
## <td headers="stat_1" class="gt_row gt_center">33 (29, 37)</td>
## <td headers="stat_2" class="gt_row gt_center">46 (43, 50)</td></tr>
##     <tr><td headers="label" class="gt_row gt_left">White Blood Cell Count</td>
## <td headers="stat_1" class="gt_row gt_center">8,800 (6,700, 10,700)</td>
```

```
## <td headers="stat_2" class="gt_row gt_center">7,500 (6,300, 9,300)</td></tr>
##   </tbody>
##
##   <tfoot class="gt_footnotes">
##     <tr>
##       <td class="gt_footnote" colspan="3"><span class="gt_footnote_marks" style="white-space:nowrap;
##     </tr>
##   </tfoot>
## </table>
## </div>
```

Citations: Rubini, L., Soundarapandian, P., & Eswaran, P. (2015). Chronic Kidney Disease [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C5G020.

Vaidya SR, Aeddula NR. Chronic Kidney Disease. [Updated 2024 Jul 31]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK535404/