

Analisis univariado

Contextuales de la persona

```
df1 %>%
  me_intervalo("edad") %>%
  me_pivot("Intervalo de edad")

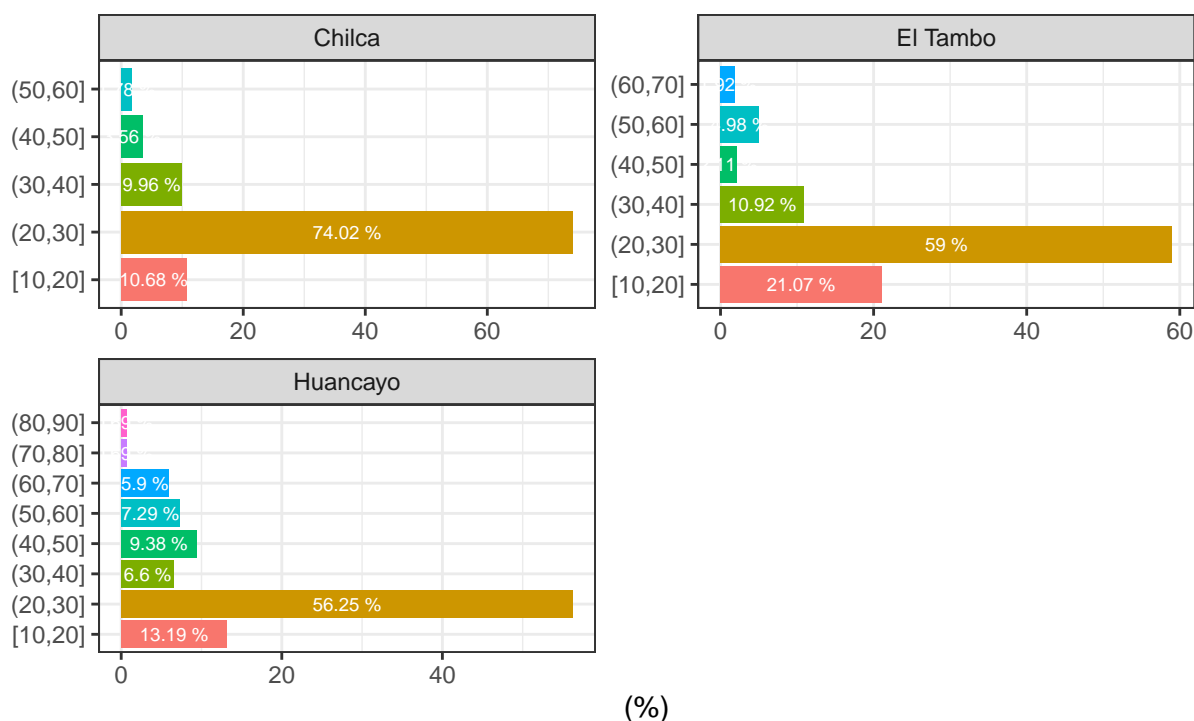
## `mutate_if()` ignored the following grouping variables:
## Column `variable`

## # A tibble: 8 x 5
## # Groups:   variable [1]
##   variable `Intervalo de edad` Chilca `El Tambo` Huancayo
##   <chr>    <fct>                <chr>  <chr>      <chr>
## 1 Edad    [10,20]                10.68 % 21.07 %   13.19 %
## 2 Edad    (20,30]                74.02 % 59 %      56.25 %
## 3 Edad    (30,40]                9.96 % 10.92 %   6.6 %
## 4 Edad    (40,50]                3.56 % 2.11 %    9.38 %
## 5 Edad    (50,60]                1.78 % 4.98 %    7.29 %
## 6 Edad    (60,70]                0 %    1.92 %    5.9 %
## 7 Edad    (70,80]                0 %    0 %       0.69 %
## 8 Edad    (80,90]                0 %    0 %       0.69 %

df1 %>%
  me_intervalo("edad") %>%
  meb_plot("Edad", "Intervalos de 10 anios", pos = "none")
```

Edad

Intervalos de 10 anios



Fuente: Encuesta realizada octubre 2020

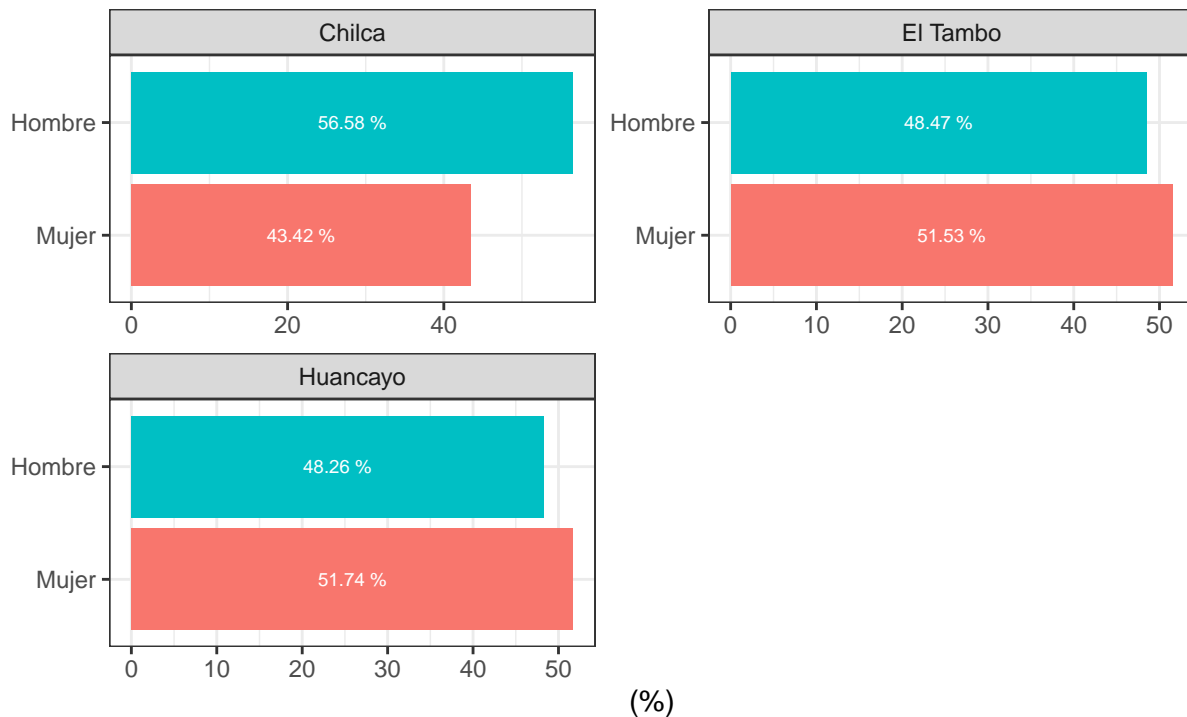
- genero (bar)

```
df1 %>%
  me_filter("genero") %>%
  me_pivot("Genero")

## # A tibble: 2 x 5
##   variable      Genero Chilca `El Tambo` Huancayo
##   <chr>          <fct> <chr>   <chr>    <chr>
## 1 Genero del encuestado Hombre 56.58 % 48.47 % 48.26 %
## 2 Genero del encuestado Mujer  43.42 % 51.53 % 51.74 %

df1 %>%
  me_filter("genero") %>%
  meb_plot("Genero", pos = "none")
```

Genero



Fuente: Encuesta realizada octubre 2020

- idiomas (tree)

```
df1 %>%
  me_filter("idiomas") %>%
  me_pivot("Numero de idiomas")

df1 %>%
  me_filter("idiomas") %>%
  filter(value != "0") %>%
  meb_plot(pos = "none")
```

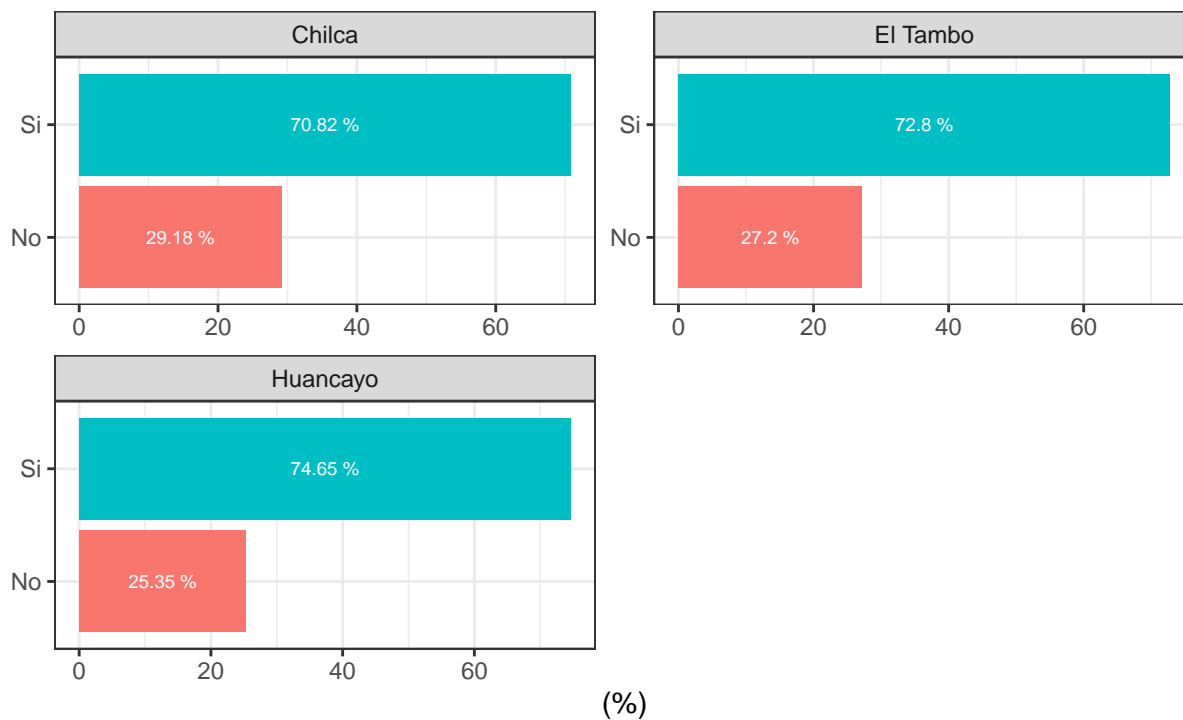
- sentimiento_serguridad (bar)

```
df1 %>%
  me_filter("sentimiento_serguridad") %>%
  me_pivot("Se siente seguro")
```

```
## # A tibble: 2 x 5
##   variable                `Se siente segur~ Chilca `El Tambo` Huancayo
##   <chr>                  <fct>          <chr>  <chr>    <chr>
## 1 Sentimiento de seguridad vivierend~ No          29.18~ 27.2 %    25.35 %
## 2 Sentimiento de seguridad vivierend~ Si          70.82~ 72.8 %    74.65 %
```

```
df1 %>%
  me_filter("sentimiento_serguridad") %>%
  meb_plot(pos = "none", "Se siente seguro viviendo en su barrio")
```

Se siente seguro viviendo en su barrio



Fuente: Encuesta realizada octubre 2020

- satisfaccion (bar)

```
df1 %>%
  me_filter("satisfaccion") %>%
  me_pivot("Sentimiento de \nsatisfaccion")
```

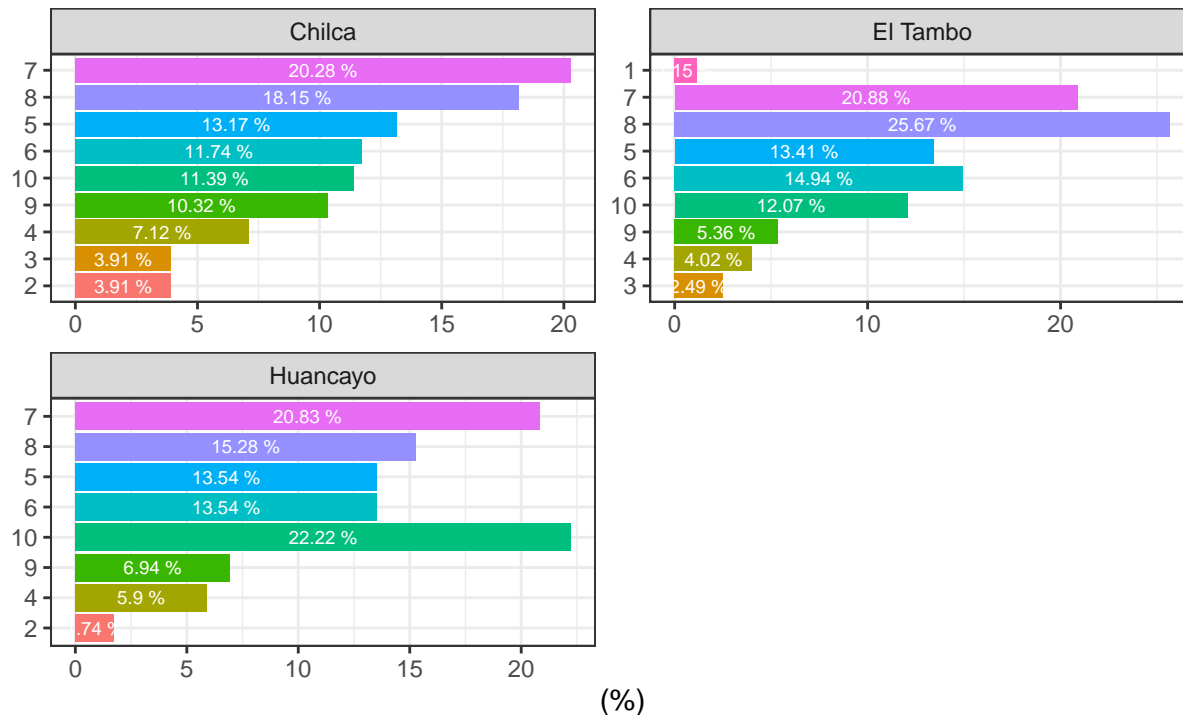
```
## # A tibble: 10 x 5
##   variable                `Sentimiento de \nsati~ Chilca `El Tambo` Huancayo
##   <chr>                  <fct>          <chr>  <chr>    <chr>
## 1 Sentimiento de satisfacci~ 10          11.39~ 12.07 %    22.22 %
## 2 Sentimiento de satisfacci~ 2           3.91 % 0 %    1.74 %
## 3 Sentimiento de satisfacci~ 3           3.91 % 2.49 %    0 %
## 4 Sentimiento de satisfacci~ 4           7.12 % 4.02 %    5.9 %
## 5 Sentimiento de satisfacci~ 5          13.17~ 13.41 %    13.54 %
```

```
## 6 Sentimiento de satisfacci~ 6 11.74~ 14.94 % 13.54 %
## 7 Sentimiento de satisfacci~ 7 20.28~ 20.88 % 20.83 %
## 8 Sentimiento de satisfacci~ 8 18.15~ 25.67 % 15.28 %
## 9 Sentimiento de satisfacci~ 9 10.32~ 5.36 % 6.94 %
## 10 Sentimiento de satisfacci~ 1 0 % 1.15 % 0 %
```

```
df1 %>%
  me_filter("satisfaccion") %>%
  meb_plot(pos = "non", "Satisfaccion", coll = 2, "1 [Poco satisfecho] - 10 [Muy satisfecho]")
```

Satisfaccion

1 [Poco satisfecho] – 10 [Muy satisfecho]



Fuente: Encuesta realizada octubre 2020

servicios

Servicios recibidos

– facet_wrap

Los servicios de los encuestados

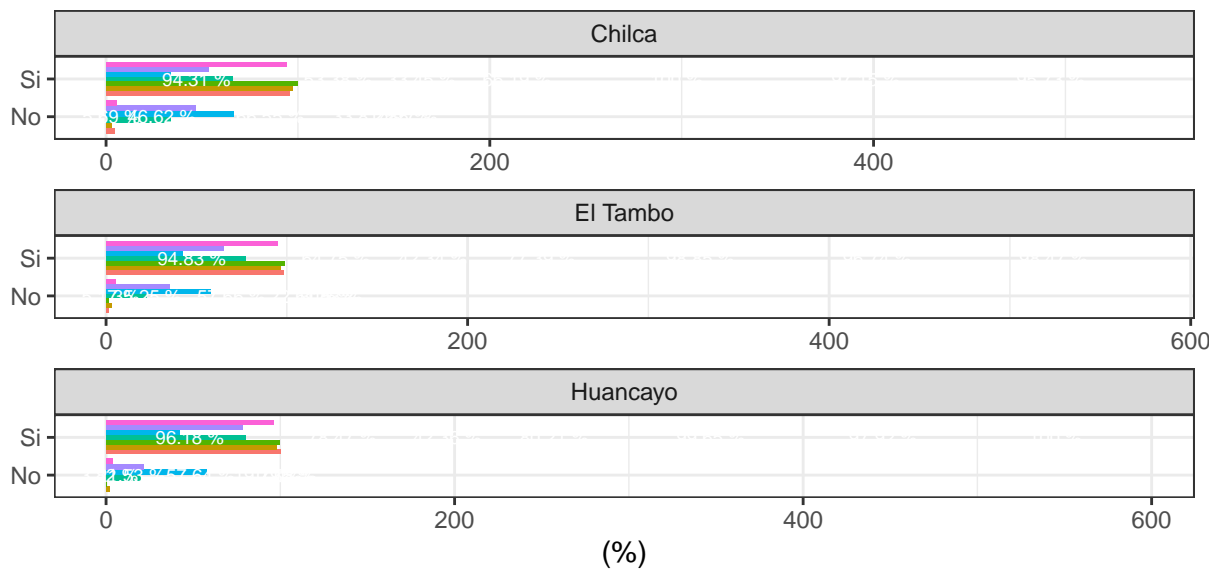
```
df1 %>%
  me_filter("ser|muni") %>%
  filter(!str_detect(variable, "[Cc]alidad|seguridad")) %>%
  arrange(1, 2) %>%
  me_pivot()
```

```
## # A tibble: 14 x 5
##   variable                                `insertar nombr~ Chilca `El Tambo` Huancayo
##   <chr>                                <fct>          <chr>  <chr>      <chr>
## 1 Acceso al servicio de agua potab~ No             4.27 % 1.53 %    0 %
## 2 Acceso al servicio de agua potab~ Si             95.73~ 98.47 % 100 %
```

## 3	Acceso al servicio de desagüe po~	No	2.85 %	3.26 %	2.08 %
## 4	Acceso al servicio de desagüe po~	Si	97.15~	96.74 %	97.92 %
## 5	Acceso al servicio de electricidad~	Si	100 %	98.85 %	99.65 %
## 6	Acceso al servicio de internet e~	No	33.81~	22.61 %	19.79 %
## 7	Acceso al servicio de internet e~	Si	66.19~	77.39 %	80.21 %
## 8	Acceso al servicio de telefonía ~	No	66.55~	57.66 %	57.64 %
## 9	Acceso al servicio de telefonía ~	Si	33.45~	42.34 %	42.36 %
## 10	La municipalidad hace mantenimie~	No	46.62~	35.25 %	21.53 %
## 11	La municipalidad hace mantenimie~	Si	53.38~	64.75 %	78.47 %
## 12	La municipalidad provee el servi~	No	5.69 %	5.17 %	3.82 %
## 13	La municipalidad provee el servi~	Si	94.31~	94.83 %	96.18 %
## 14	Acceso al servicio de electricidad~	No	0 %	1.15 %	0.35 %

```
df1 %>%
  me_filter("ser|muni") %>%
  filter(!str_detect(variable, "[Cc]alidad|seguridad")) %>%
  meb_plot(coll = 1) +
  aes(fill = variable)
```

titulo



■ Acceso al servicio de electricidad en la vivienda
 ■ Acceso al servicio de telefonía fija en la vivienda
■ Acceso al servicio de internet en la vivienda
 ■ La municipalidad hace mantenimiento de los parques

Fuente: Encuesta realizada octubre 2020

Calidad de los servicios recibidos

- filtrar por los que reciben
- ~facet_grid(districto)

```
calidad <- df1 %>%
  filter(str_detect(name, "cal_")) %>%
  group_by(Districto, variable) %>%
  count(value) %>%
```

```
filter(value != 0) %>%
mutate(por = round(n / sum(n)*100, 2),
       porcentaje = paste0(por, "%")) %>%
ungroup()
```

calidad

```
## # A tibble: 119 x 6
##   Distrito variable          value      n   por porcentaje
##   <chr>      <chr>          <chr> <int> <dbl> <chr>
## 1 Chilca    Calidad del mantenimiento de parques~ Bueno      61 23.7 23.74%
## 2 Chilca    Calidad del mantenimiento de parques~ Malo       53 20.6 20.62%
## 3 Chilca    Calidad del mantenimiento de parques~ Muy bu~    3  1.17 1.17%
## 4 Chilca    Calidad del mantenimiento de parques~ Muy ma~   18  7    7%
## 5 Chilca    Calidad del mantenimiento de parques~ Regular  122 47.5 47.47%
## 6 Chilca    Calidad del servicio de agua          Bueno  114 43.0 43.02%
## 7 Chilca    Calidad del servicio de agua          Malo    14  5.28 5.28%
## 8 Chilca    Calidad del servicio de agua          Muy bu~  41 15.5 15.47%
## 9 Chilca    Calidad del servicio de agua          Muy ma~  19  7.17 7.17%
## 10 Chilca   Calidad del servicio de agua          Regular  77 29.1 29.06%
## # ... with 109 more rows
```

```
calidad %>%
  select(!c(n, por)) %>%
  mutate(porcentaje = replace_na(porcentaje, "0%")) %>%
  #group_by()
  pivot_wider( names_from = Distrito, values_from = porcentaje)
```

```
## # A tibble: 40 x 5
##   variable          value Chilca `El Tambo` Huancayo
##   <chr>          <chr> <chr> <chr> <chr>
## 1 Calidad del mantenimiento de parques y j~ Bueno 23.74% 28.23% 34.29%
## 2 Calidad del mantenimiento de parques y j~ Malo 20.62% 18.29% 17.86%
## 3 Calidad del mantenimiento de parques y j~ Muy bue~ 1.17% 7.95% 9.64%
## 4 Calidad del mantenimiento de parques y j~ Muy malo 7% 5.37% 3.57%
## 5 Calidad del mantenimiento de parques y j~ Regular 47.47% 40.16% 34.64%
## 6 Calidad del servicio de agua          Bueno 43.02% 52.6% 41.75%
## 7 Calidad del servicio de agua          Malo 5.28% 5.2% 5.96%
## 8 Calidad del servicio de agua          Muy bue~ 15.47% 12.52% 16.14%
## 9 Calidad del servicio de agua          Muy malo 7.17% 1.93% 3.51%
## 10 Calidad del servicio de agua          Regular 29.06% 27.75% 32.63%
## # ... with 30 more rows
```

Tiempo

bar chart ramas o tal vez

– correguir la base de datos si hay tiempo’, mejor es un case)when

Vivienda

las variables de metros cuadrados precios de los mismos tienen valores anormales en mi opinion

```
vivien <- df1 %>%
  filter(str_detect(variable, "Numero")) %>%
```

```

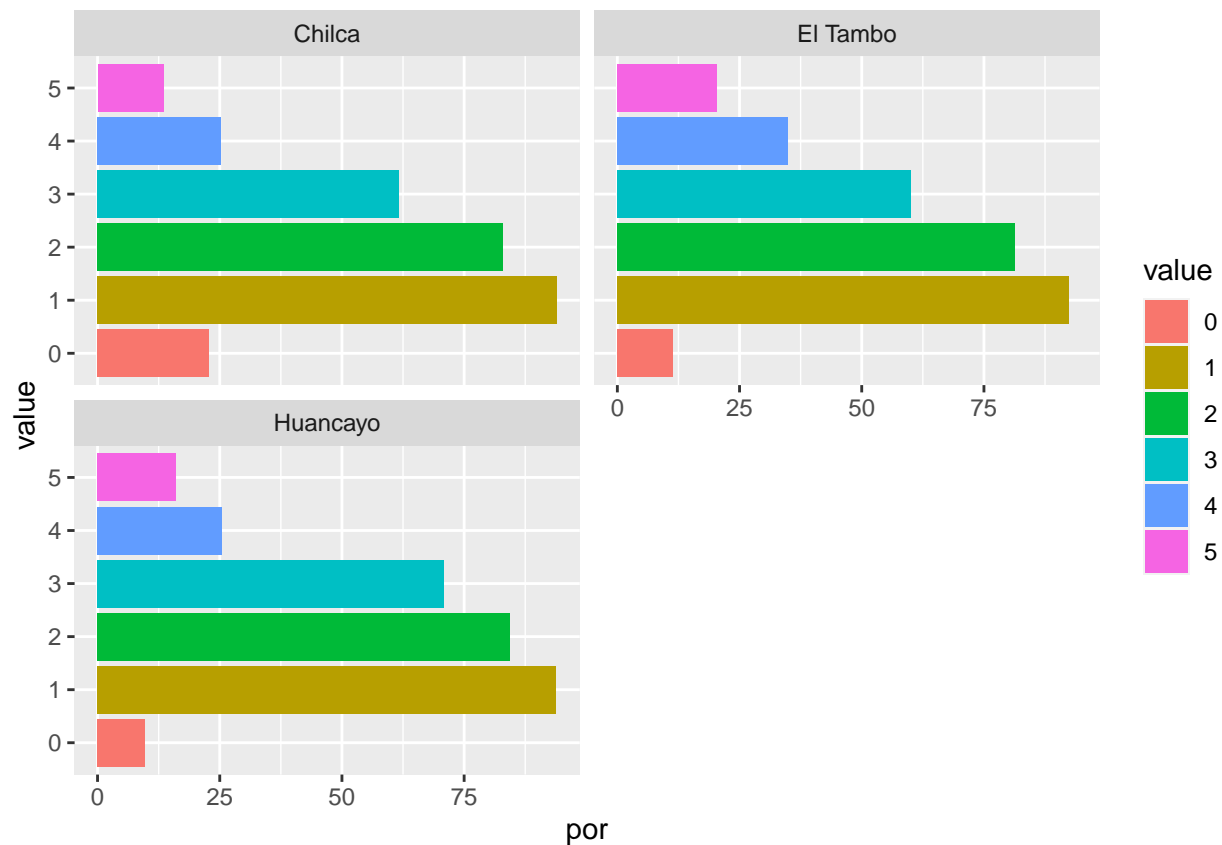
filter(str_detect(name, "vivienda")) %>%
mutate(value = parse_number(value)) %>%
group_by(Distrito, variable) %>%
count(value) %>%
#group_by(Distrito, variable) %>%
mutate(por = round(n / sum(n) * 100, 2),
       por1 = paste0(por, "%"))

vivien %>%
  select(!c(n, por)) %>%
  pivot_wider(names_from = value, values_from = por1)

## # A tibble: 9 x 8
## # Groups:   Distrito, variable [9]
##   Distrito variable      `0`      `1`      `2`      `3`      `4`      `5`
##   <chr>      <chr>      <chr> <chr> <chr> <chr> <chr>
## 1 Chilca    Numero de banios de la vivi~ 10.68% 39.15% 24.91% 18.86% 6.41% <NA>
## 2 Chilca    Numero de cuartos de la viv~ 1.78% 38.08% 12.81% 23.49% 11.39% 12.4~
## 3 Chilca    Numero de pisos de la vivie~ 10.32% 16.73% 45.2% 19.22% 7.47% 1.07%
## 4 El Tambo  Numero de banios de la vivi~ 4.21% 31.61% 32.18% 19.54% 8.05% 4.41%
## 5 El Tambo  Numero de cuartos de la viv~ <NA> 39.08% 14.37% 18.58% 17.24% 10.7~
## 6 El Tambo  Numero de pisos de la vivie~ 7.09% 21.65% 34.67% 21.84% 9.58% 5.17%
## 7 Huancayo  Numero de banios de la vivi~ 6.25% 30.21% 34.72% 17.71% 9.38% 1.74%
## 8 Huancayo  Numero de cuartos de la viv~ 1.04% 37.85% 7.99% 32.29% 9.72% 11.1~
## 9 Huancayo  Numero de pisos de la vivie~ 2.43% 25.69% 41.67% 20.83% 6.25% 3.12%

vivien %>%
  mutate(value = factor(value)) %>%
  ggplot() +
  geom_col() +
  aes(value, por, fill = value) +
  facet_wrap(~Distrito, ncol = 2) +
  coord_flip()

```

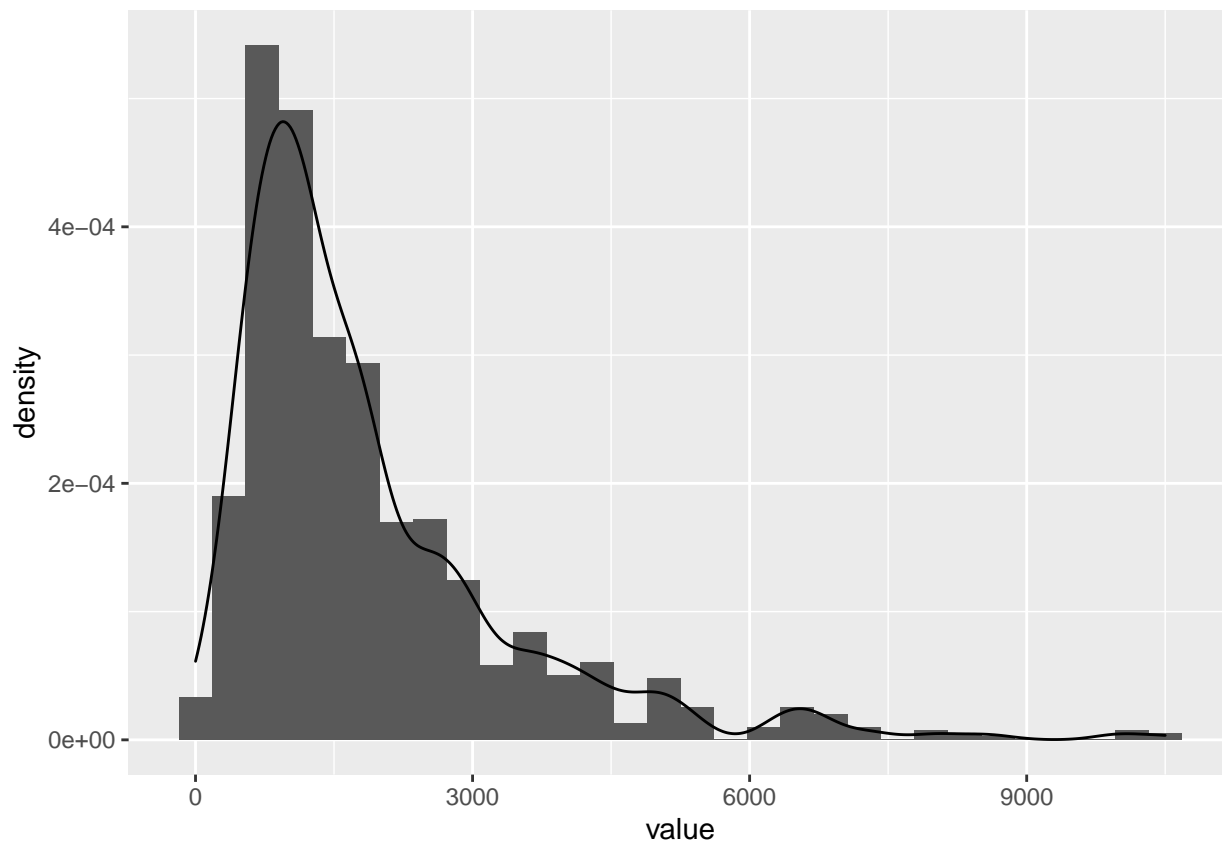


Gasto (gasto total)

```
gasto <- df1 %>%
  filter(str_detect(name, "gasto")) %>%
  select(!c(id, name)) %>%
  mutate(value = parse_number(value))

gasto %>%
  ggplot() +
  aes(value) +
  geom_histogram(aes(y = ..density..)) +
  geom_density()
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.



Al rededores

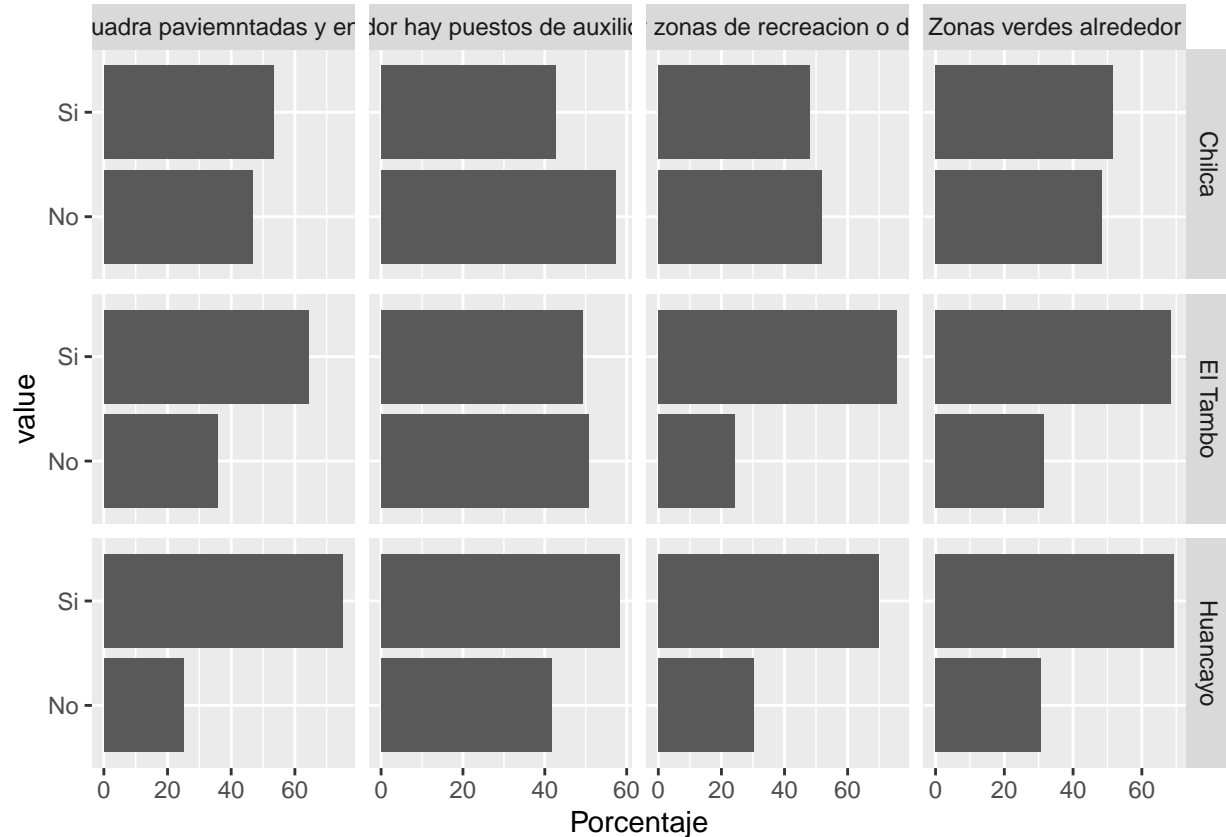
```
alr <- df1 %>%
  filter(str_detect(name, "alr")) %>%
  select(!c(id, name)) %>%
  group_by(Distrito, variable) %>%
  count(value) %>%
  mutate(Porcentaje = round(n / sum(n) * 100, 2),
         por1 = paste0(Porcentaje, "%"))

alr %>%
  select(!c(n, Porcentaje)) %>%
  pivot_wider(names_from = value, values_from = por1)
```

```
## # A tibble: 12 x 4
## # Groups:   Distrito, variable [12]
##   Distrito variable                                No    Si
##   <chr>    <chr>                                <chr> <chr>
## 1 Chilca Acceso a la cuadra paviemntadas y en buen estado 46.62% 53.38%
## 2 Chilca Al rededor hay puestos de auxilio rapido 57.3% 42.7%
## 3 Chilca Paques y zonas de recreacion o deportivas 51.96% 48.04%
## 4 Chilca Zonas verdes alrededor 48.4% 51.6%
## 5 El Tambo Acceso a la cuadra paviemntadas y en buen estado 35.63% 64.37%
## 6 El Tambo Al rededor hay puestos de auxilio rapido 50.77% 49.23%
## 7 El Tambo Paques y zonas de recreacion o deportivas 24.33% 75.67%
## 8 El Tambo Zonas verdes alrededor 31.61% 68.39%
## 9 Huancayo Acceso a la cuadra paviemntadas y en buen estado 25% 75%
```

```
## 10 Huancayo Al rededor hay puestos de auxilio rapido 41.67% 58.33%
## 11 Huancayo Paques y zonas de recreacion o deportivas 30.21% 69.79%
## 12 Huancayo Zonas verdes alrededor 30.56% 69.44%
```

```
alr %>%
  ungroup() %>%
  ggplot() +
  aes(value, Porcentaje) +
  geom_col() +
  facet_grid(Distrito~variable, scales = 'free') +
  coord_flip()
```



Familia

```
df1 %>%
  vari(var = "jefe_ocu") %>%
  t_vari
df1 %>%
  vari(var = "jefe_ocu") %>%
  b_plot()
```

- jefe_ocu
- hogar_ingresos_cuantos

```
df1 %>%
  vari(var = "hogar_ingresos_cuantos") %>%
  t_vari
df1 %>%
```

```

vari(var = "hogar_ingresos_cuantos") %>%
b_plot()

fami <- df1 %>%
  filter(str_detect(name, "f_")) %>%
  select(!c(id, name)) %>%
  group_by(Distrito, variable) %>%
  count(value) %>%
  mutate(por = round(n / sum(n) * 100, 2),
         por1 = paste0(por, "%")) %>%
  ungroup()
fami %>%
  select(!c(n, por)) %>%
  pivot_wider(names_from = value, values_from = por1)

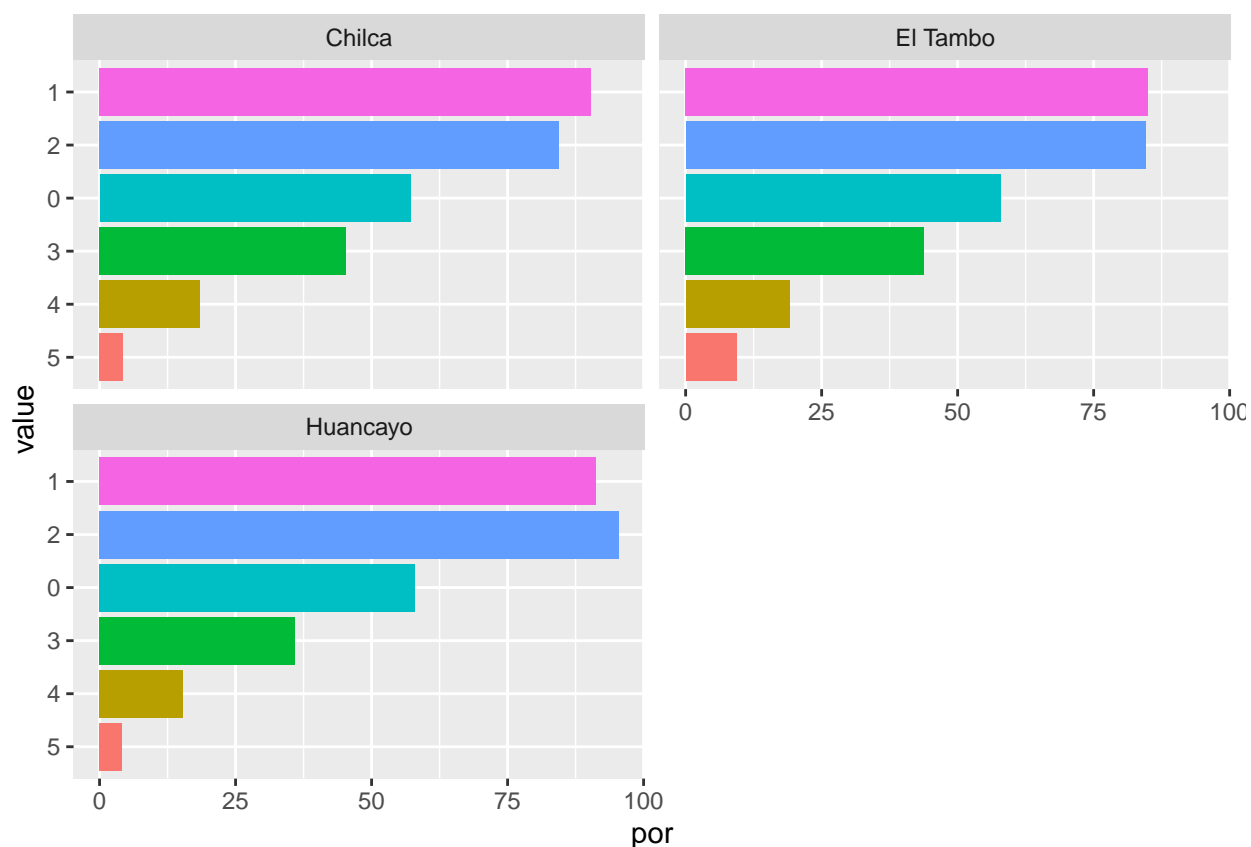
## # A tibble: 9 x 8
##   Distrito variable      `0`      `1`      `2`      `3`      `4`      `5`
##   <chr>      <chr>      <chr> <chr> <chr> <chr> <chr>
## 1 Chilca  Cantidad de menores a 18 anios 44.84% 23.49% 19.93% 9.25% 1.42% 1.07%
## 2 Chilca  Hombres mayores a 17 anios 5.69% 33.1% 32.38% 18.1~ 9.61% 1.07%
## 3 Chilca  Mujeres mayores a 17 anioes 6.76% 33.81% 32.03% 17.7~ 7.47% 2.14%
## 4 El Tambo Cantidad de menores a 18 anios 40.8% 28.16% 17.05% 6.7% 3.83% 3.45%
## 5 El Tambo Hombres mayores a 17 anios 9% 30.08% 31.42% 18.5~ 7.85% 3.07%
## 6 El Tambo Mujeres mayores a 17 anioes 8.24% 26.82% 36.21% 18.3~ 7.47% 2.87%
## 7 Huancayo Cantidad de menores a 18 anios 36.11% 29.51% 21.88% 5.21% 5.21% 2.08%
## 8 Huancayo Hombres mayores a 17 anios 14.58% 30.21% 35.07% 15.2~ 3.82% 1.04%
## 9 Huancayo Mujeres mayores a 17 anioes 7.29% 31.6% 38.54% 15.2~ 6.25% 1.04%

## eliminar la parte de composicion familiar esto puede ir la descripcion de l titiulo de la figura
fami

## # A tibble: 54 x 6
##   Distrito variable      value      n      por por1
##   <chr>      <chr>      <chr> <int> <dbl> <chr>
## 1 Chilca  Cantidad de menores a 18 anios 0      126 44.8 44.84%
## 2 Chilca  Cantidad de menores a 18 anios 1       66 23.5 23.49%
## 3 Chilca  Cantidad de menores a 18 anios 2       56 19.9 19.93%
## 4 Chilca  Cantidad de menores a 18 anios 3       26 9.25 9.25%
## 5 Chilca  Cantidad de menores a 18 anios 4        4 1.42 1.42%
## 6 Chilca  Cantidad de menores a 18 anios 5        3 1.07 1.07%
## 7 Chilca  Hombres mayores a 17 anios 0       16 5.69 5.69%
## 8 Chilca  Hombres mayores a 17 anios 1       93 33.1 33.1%
## 9 Chilca  Hombres mayores a 17 anios 2       91 32.4 32.38%
## 10 Chilca Hombres mayores a 17 anios 3       51 18.2 18.15%
## # ... with 44 more rows

fami %>%
  mutate(value = reorder(value, n)) %>%
  ggplot() +
  aes(x = value, y = por, fill = value) +
  geom_col() +
  facet_wrap(~Distrito, ncol = 2) +
  coord_flip() +
  theme(legend.position = "none")

```



Covid

El efecto del covid en la poblacion del Distrito metropolitano de Huancayo de los cuales los distritos de Chilca, El Tambo tuvieron diferentes numeros de contagios, por *por ejemplo*, mientras que la encuesta realizada el entorno contextual muestra que la variable relacionada con los casos positivos (ver siguiente tabla).

El numero de contagios como proporccion del numero de contagios sobre el total de encuestados muestra que Chilca, tuvo 14.9% de los encuestados tuvieron COVID 19, en El tambo el porcentaje de contagios fue del 12.3% mientras que en el Distrito de Huancayo solo el 6.9% fueron casos positivos de COVID 19. Esto se puede visualizar mejor en el grafico subsiguiente donde se distingue

```
df1 %>%
  vari(var = "covid_positivo") %>%
  t_vari()
df1 %>%
  vari(var = "covid_positivo") %>%
  b_plot()
```

Mientras que el porcentaje de los encuestados que se hicieron la prueba para saber si tuvieron o no COVID 19 tuvo un comportamiento diferente al anterior, ya que en el distrito de Chilca el 38.1% de los encuestados afirmaron que se hicieron una prueba para saber si tenian o no COVID 19, mientras que el 61.9% no se hicieron ni recibieron ninguna prueba para detectar COVID 19, mientras que en distrito de El Tambo el porcentaje de los encuestados que se hicieron alguna prueba para saber si tuvieron COVID 19 solo el 33.1% afirmaron que si se hicieron en contraparte los que no se hicieron ninguna ni tampoco recibieron ninguna de las pruebas fueron el 66.9%. Por ultimo en el Distrito de Huancayo es el distrito del cual menos encuestados se hicieron la prueba para COVID 19 siendo el porcentaje de 29.9%, esto se puede visualizar de mejor manera en la subsiguiente figura.

```
df1 %>%
  vari(var = "covid_prueba") %>%
  t_vari
df1 %>%
  vari(var = "covid_prueba") %>%
  b_plot()
```

Por otra parte el porcentaje de los encuestados que presentaron algunos sintomas de los caules son caracteristicos y que mediante estudios avalados por la OMS son propios del COVID 19, son tos congestion o secrecion nasal, fatiga, perdida de sentido del gusto o del olfato y dolor de cabeza o en el pecho, y mediante la encuesta en el distrito de Chilca tuvo el menor de casos con sintomas siendo este solo el 22.8%, en contraparte los que no presentaron sintomas fueron el 77.2%, mientras que el distrito de El Tambo el 23.6% presentaron estos sintomas mientras que el 76.4% no presentaron ninguno de estos sintomas. Por ultimo en el distrito de Huancayo el 25.7% el mayor a nivel distrital mostraron estos sintomas. mientras que el 74.3% de los encuestados no presentaron ninguno de los sintomas del COVID 19 (ver siguiente tabla y figura).

```
df1 %>%
  vari(var = "covid_sintomas") %>%
  t_vari
df1 %>%
  vari(var = "covid_sintomas") %>%
  b_plot()
```

Por otra parte se sabe que el virus necesita de un portador humano para poder transmitirse algunos son asintomaticos (no presentan sintomas) y otros si presentan sintomas, y para poder recabar esta informacion, se tuvieron 2 preguntas, la primera era si se conocia a alguien de la comunidad que estuviese enfermo de coronavirus (estuvo enfermo, tos, dificultad para respirar, etc.), y la segunda si hubo contacto con alguna persona la cual presentaba sintomas propios del coronavirus, por distritos se tiene que para los conocidos en los distritos de Chilca, El Tambo y Huancayo fueron de 69.8%, 68.4%, y 60.8% respectivamente, mientras que los contactos registrados con alguna persona sospechoza de COVID 19 fueron casi la tercera parte, siendo estas para los distritos de Chilca, El Tambo y Huancayo de 27.4% y 28%, y 29.5%, mostrando la preocupacion de tomar algo de distancia con las personas que presenta o presentaron sintomas.

```
df1 %>%
  vari(var = "covid_conocido_sintomas") %>%
  t_vari
df1 %>%
  vari(var = "covid_conocido_sintomas") %>%
  b_plot()
df1 %>%
  vari(var = "covid_contacto_sintomas") %>%
  t_vari
df1 %>%
  vari(var = "covid_contacto_sintomas") %>%
  b_plot()
```

Asimismo existen gastos destinados principalmente para prevenir y tratar el COVID 19, para prevenir se tiene la higiene como principal recurso, este puede comprender al uso de agua y jabon, desinfectantes de manos, mascarillas, protectores faciales entre otros, mientras que para el tratamiento, puede ser de la persona como tambien de la familia, y este comprende la adquisicion de medicamentos, gastos medicos, balones de oxigeno, los encuestados mostraron el siguiente comportamiento.

Esta parte crear una funcion de histograma, y facet -1 col esto se repite para lo anterior,
pensado para estas variables

- covid __ desempleo
- covid__ drepsion
- covid_conflictos

Analisis bivariado

Modelo econométrico