```
library(tidyverse)
library(ggplot2)
library(scales)
sba <- read_csv("https://uofi.box.com/shared/static/vi37omgitiaa2yyplrom779qvwk1g14x.csv",
    col_types = cols(ApprovalDate = col_date(format = "%d-%b-%y"),
                     BalanceGross = col_number(), ChgOffDate = col_date(format = "%d-%b-%y"),
                     ChgOffPrinGr = col_number(), DisbursementDate = col_date(format ="%d-%b-%y"),
                     DisbursementGross = col_number(),
                     ApprovalFY = col_integer(),
                     GrAppv = col_number(), SBA_Appv = col_number()))
colnames(sba) <- tolower(colnames(sba))
```

```
area <- factor(sba$urbanrural,c(0,1,2),labels=c("Undefined","Urban","Rural"))
meow <- table(area,sba$mis_status)
```

```
group_by(sba,newexist)
```

```
## Warning: One or more parsing issues, see `problems()` for details
```

```
## # A tibble: 899,164 x 27
## # Groups:   newexist [4]
##    loannr_chkdgt name    city   state   zip bank   bankstate  naics approvaldate
##            <dbl> <chr>   <chr>  <chr> <dbl> <chr>  <chr>      <dbl> <date>
##  1    1000014003 ABC HO~ EVANS~ IN    47711 FIFTH~ OH        451120 1997-02-28
##  2    1000024006 LANDMA~ NEW P~ IN    46526 1ST S~ IN        722410 1997-02-28
##  3    1000034009 WHITLO~ BLOOM~ IN    47401 GRANT~ IN        621210 1997-02-28
##  4    1000044001 BIG BU~ BROKE~ OK    74012 1ST N~ OK             0 1997-02-28
##  5    1000054004 ANASTA~ ORLAN~ FL    32801 FLORI~ FL             0 1997-02-28
##  6    1000084002 B&T SC~ PLAIN~ CT     6062 TD BA~ DE        332721 1997-02-28
##  7    1000093009 MIDDLE~ UNION  NJ     7083 WELLS~ SD             0 NA
##  8    1000094005 WEAVER~ SUMME~ FL    34491 REGIO~ AL        811118 1997-02-28
##  9    1000104006 TURTLE~ PORT ~ FL    32456 CENTE~ FL        721310 1997-02-28
## 10    1000124001 INTEXT~ GLAST~ CT     6073 WEBST~ CT             0 1997-02-28
## # ... with 899,154 more rows, and 18 more variables: approvalfy <int>,
## #   term <dbl>, noemp <dbl>, newexist <dbl>, createjob <dbl>,
## #   retainedjob <dbl>, franchisecode <dbl>, urbanrural <dbl>, revlinecr <chr>,
## #   lowdoc <chr>, chgoffdate <date>, disbursementdate <date>,
## #   disbursementgross <dbl>, balancegross <dbl>, mis_status <chr>,
## #   chgoffpringr <dbl>, grappv <dbl>, sba_appv <dbl>
```

```
sba$estb <- ifelse(sba$newexist==2,"No","Yes")
sba$ms_new <- ifelse(sba$mis_status=="P I F","Paid in Full","Defaulted")
#table1 <- table(sba$estb,area,sba$ms_new);table1
#ftable(round(prop.table(table1)*100, 2))
#deflab <- c(24.49,1.70,8.94,0.86,28.71,9.12,10.79,3.68,6.36,1.46,3.15,0.73)
```

```
theme_set(theme_minimal())
ggplot(data=sba,aes(x=factor(estb),y =(..count..)/sum(..count..),fill=factor(ms_new)))+
  geom_bar(width=.75,na.rm=TRUE)+
  scale_y_continuous(labels=percent)+
  facet_wrap(~factor(area),ncol=1)+
```

```
labs(x="Established Business?",y="Percentage",title="Loan Result by Area & Establishment of Businesses
guides(fill=guide_legend(title="Loan Status"))+
coord_flip()
```