

Quiz 6

Name:

titanic data

For all of the questions today, you will be using data for passengers from the Titanic. Here is the start of this dataset, called `titanic`:

```
titanic <- read.csv("../data/titanic.csv", as.is = TRUE) %>%
  select(Survived, Pclass, Age, Name, Sex, Fare) %>%
  mutate(Pclass = as.factor(Pclass))
head(titanic)
```

```
##   Survived Pclass Age                               Name
## 1         0      3  22                               Braund, Mr. Owen Harris
## 2         1      1  38 Cumings, Mrs. John Bradley (Florence Briggs Thayer)
## 3         1      3  26                               Heikkinen, Miss. Laina
## 4         1      1  35 Futrelle, Mrs. Jacques Heath (Lily May Peel)
## 5         0      3  35                               Allen, Mr. William Henry
## 6         0      3  NA                               Moran, Mr. James
##      Sex   Fare
## 1   male 7.2500
## 2 female 71.2833
## 3 female  7.9250
## 4 female 53.1000
## 5   male  8.0500
## 6   male  8.4583
```

1.-12.

Match the number at the end of each line to the letter of the output you would have if you ran all code up to that point. For all dataframes, only the first few lines are shown. Put the letter answer next to the number given in the code.

```
titanic$agecat[titanic$Age < 15] <- "Under 15"           # 1.
titanic$agecat[titanic$Age >= 15 &
  titanic$Age <= 50] <- "15--50"                       # 2.
titanic$agecat[titanic$Age > 50] <- "Over 50"
titanic$agecat <- factor(titanic$agecat,
  levels = c("Under 15",
             "15--50",
             "Over 50"))

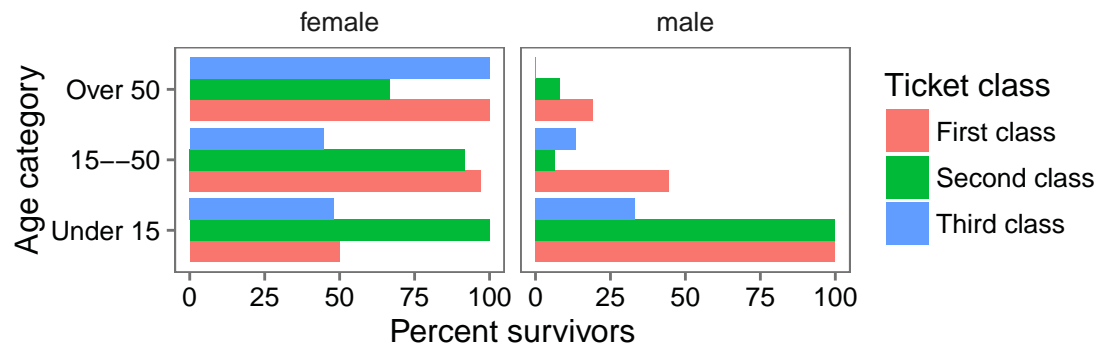
toplot_4 <- filter(titanic, !is.na(Age)) %>%             # 3.
  select(Survived, Pclass, agecat, Sex) %>%             # 4.
  group_by(Pclass, agecat, Sex) %>%
  summarize(N = n(),
    survivors = sum(Survived == 1),
    perc_survived = 100 * survivors / N) %>% #5.
  ungroup() %>%
```

```

mutate(Pclass = factor(Pclass, levels = c(1, 2, 3),
                        labels = c("First class", "Second class",
                                   "Third class")))) # 6.
plot_4 <- ggplot(topplot_4, aes(x = agecat, y = perc_survived, fill = Pclass)) +
  geom_bar(stat = "identity", position = "dodge") + # 7.
  coord_flip() + # 8.
  ylab("Percent survivors") + xlab("Age category") + # 9.
  theme_few() + # 10.
  scale_fill_discrete(name = "Ticket class") + # 11.
  facet_grid(. ~ Sex) # 12.

```

a.



b.

| ## | Survived | Pclass | Age | Name |
|------|----------|--------|-----|--|
| ## 1 | 0 | 3 | 22 | Braund, Mr. Owen Harris |
| ## 2 | 1 | 1 | 38 | Cummings, Mrs. John Bradley (Florence Briggs Thayer) |
| ## 3 | 1 | 3 | 26 | Heikkinen, Miss. Laina |
| ## 4 | 1 | 1 | 35 | Futrelle, Mrs. Jacques Heath (Lily May Peel) |
| ## 5 | 0 | 3 | 35 | Allen, Mr. William Henry |
| ## 6 | 0 | 3 | NA | Moran, Mr. James |

| ## | Sex | Fare | agecat |
|------|--------|---------|--------|
| ## 1 | male | 7.2500 | <NA> |
| ## 2 | female | 71.2833 | <NA> |
| ## 3 | female | 7.9250 | <NA> |
| ## 4 | female | 53.1000 | <NA> |
| ## 5 | male | 8.0500 | <NA> |
| ## 6 | male | 8.4583 | <NA> |

c.

```

## Source: local data frame [6 x 6]
## Groups: Pclass, agecat [3]
##
##   Pclass  agecat  Sex    N survivors perc_survived
##   (fctr)  (fctr)  (chr) (int)      (int)          (dbl)
## 1     1 Under 15 female    2         1         50.00000
## 2     1 Under 15  male    3         3        100.00000
## 3     1  15--50 female   70        68         97.14286

```

```
## 4      1  15--50  male   72      32      44.44444
## 5      1  Over 50 female   13      13     100.00000
## 6      1  Over 50  male   26       5     19.23077
```

d.

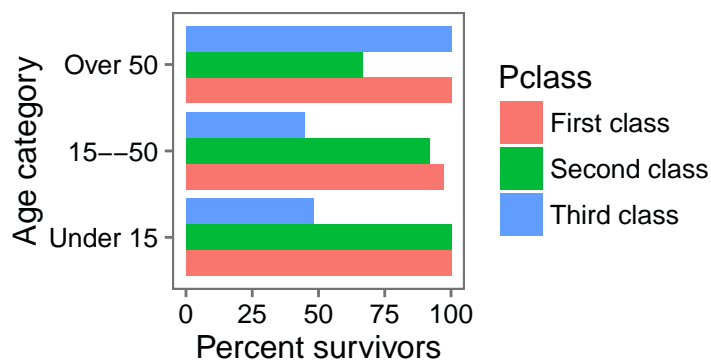
```
##   Survived Pclass Age                                     Name
## 1         0      3  22                                     Braund, Mr. Owen Harris
## 2         1      1  38 Cumings, Mrs. John Bradley (Florence Briggs Thayer)
## 3         1      3  26                                     Heikkinen, Miss. Laina
## 4         1      1  35      Futrelle, Mrs. Jacques Heath (Lily May Peel)
## 5         0      3  35                                     Allen, Mr. William Henry
## 6         0      1  54      McCarthy, Mr. Timothy J

##      Sex   Fare  agecat
## 1   male  7.2500  15--50
## 2 female 71.2833  15--50
## 3 female  7.9250  15--50
## 4 female 53.1000  15--50
## 5   male  8.0500  15--50
## 6   male 51.8625  Over 50
```

e.

```
##   Survived Pclass  agecat   Sex
## 1         0      3  15--50  male
## 2         1      1  15--50 female
## 3         1      3  15--50 female
## 4         1      1  15--50 female
## 5         0      3  15--50  male
## 6         0      1  Over 50  male
```

f.



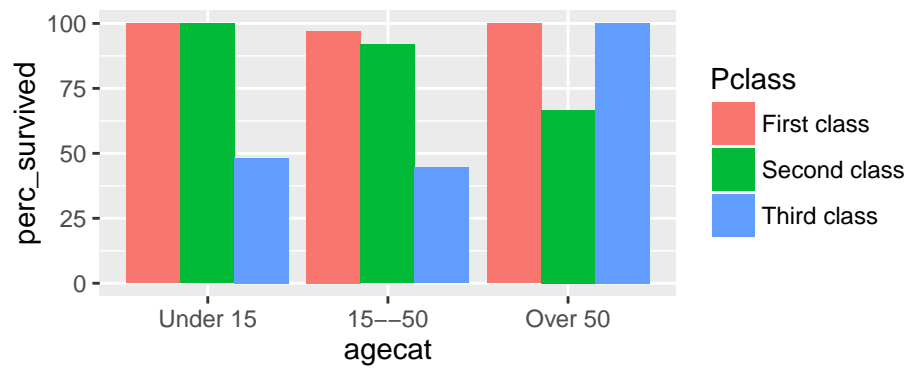
g.

```
##   Survived Pclass Age                                     Name
## 1         0      3  22                                     Braund, Mr. Owen Harris
## 2         1      1  38 Cumings, Mrs. John Bradley (Florence Briggs Thayer)
## 3         1      3  26                                     Heikkinen, Miss. Laina
## 4         1      1  35      Futrelle, Mrs. Jacques Heath (Lily May Peel)
```

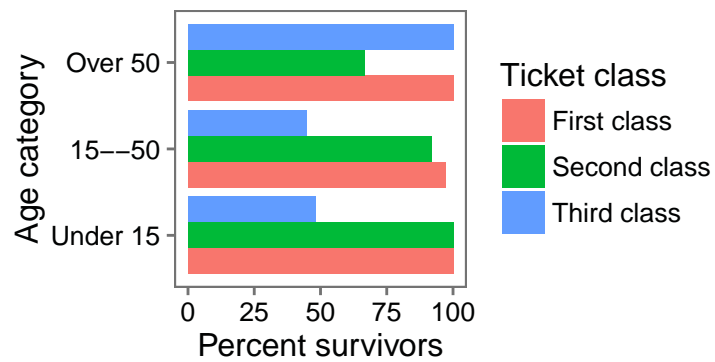
```
## 5      0      3 35
## 6      0      3 NA
##      Sex      Fare agecat
## 1   male  7.2500 15--50
## 2 female 71.2833 15--50
## 3 female  7.9250 15--50
## 4 female 53.1000 15--50
## 5   male  8.0500 15--50
## 6   male  8.4583  <NA>
```

```
Allen, Mr. William Henry
Moran, Mr. James
```

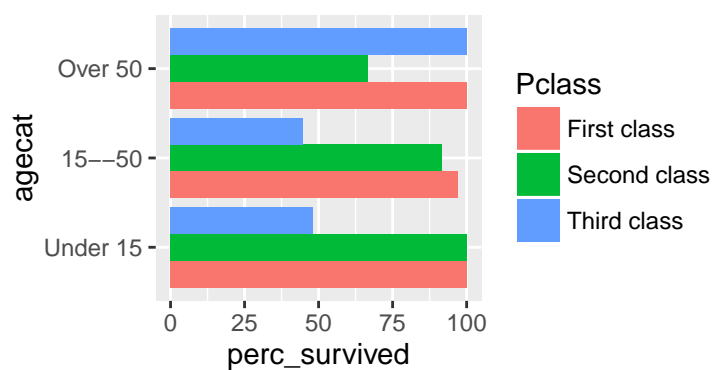
h.



i.



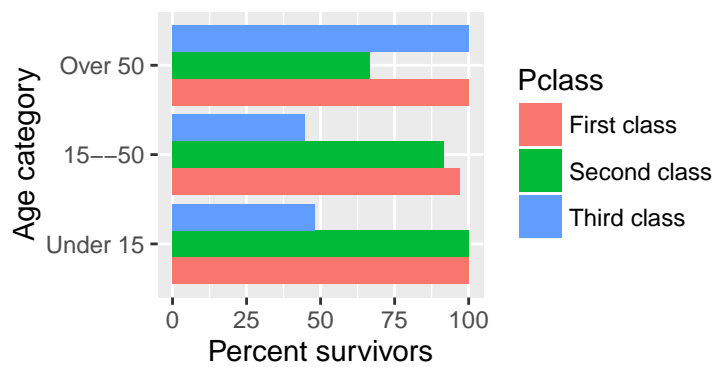
j.



k.

```
## Source: local data frame [6 x 6]
##
##      Pclass  agecat  Sex    N survivors perc_survived
##      (fctr)  (fctr) (chr) (int)      (int)          (dbl)
## 1 First class Under 15 female    2         1      50.00000
## 2 First class Under 15  male    3         3     100.00000
## 3 First class  15--50 female   70        68     97.14286
## 4 First class  15--50  male   72        32     44.44444
## 5 First class  Over 50 female   13        13     100.00000
## 6 First class  Over 50  male   26         5     19.23077
```

l.



Answers: 1. b., 2. g., 3. d., 4. e., 5. c., 6. k., 7. h., 8. j., 9. l., 10. f., 11. i., 12. a.