

# Challenge 2 Answer Key

## 1. Change sex and ulcer to categorical variables using mutate. (5 points)

```
Melanoma <- Melanoma %>%  
  mutate(sex = factor(sex, labels = c("Female", "Male")), ulcer = factor(ulcer))
```

1 Point: Correct labels for sex

1 Point: dplyr used

3 Points: sex and ulcer changed to factors

## 2. Add died variable to the dataset. (5 points)

```
Melanoma <- Melanoma %>% mutate(died = ifelse(status == 1, 1, 0))
```

1 Point: dplyr used

4 Points: died added correctly

## 3. Create a new table using the pipe command. (5 points)

```
Melanoma %>%  
  group_by(ulcer) %>%  
  filter(age > 65) %>%  
  summarise(n = n(), age = mean(age), thickness = mean(thickness), died = mean(died)) %>%  
  arrange(desc(n))
```

2 Points: group\_by variable changed to ulcer

1 Point: Correct age range selected

2 Points: Table created and interpreted correctly

## 4. Create a new subset of the data using dplyr, then fit a logistic regression model to it using the glm function. (5 points)

```
Melanoma2 <- Melanoma %>% dplyr::select(sex, age, thickness, ulcer, died)  
mod <- glm(died~., family=binomial(link='logit'), data = Melanoma2)  
summary(mod)
```

1 Point: select used to create new dataframe

2 Points: glm function used to fit model

2 Points: Model coefficients interpreted correctly

*If warnings or messages are included in the RMarkdown output, deduct one point from the final score.*