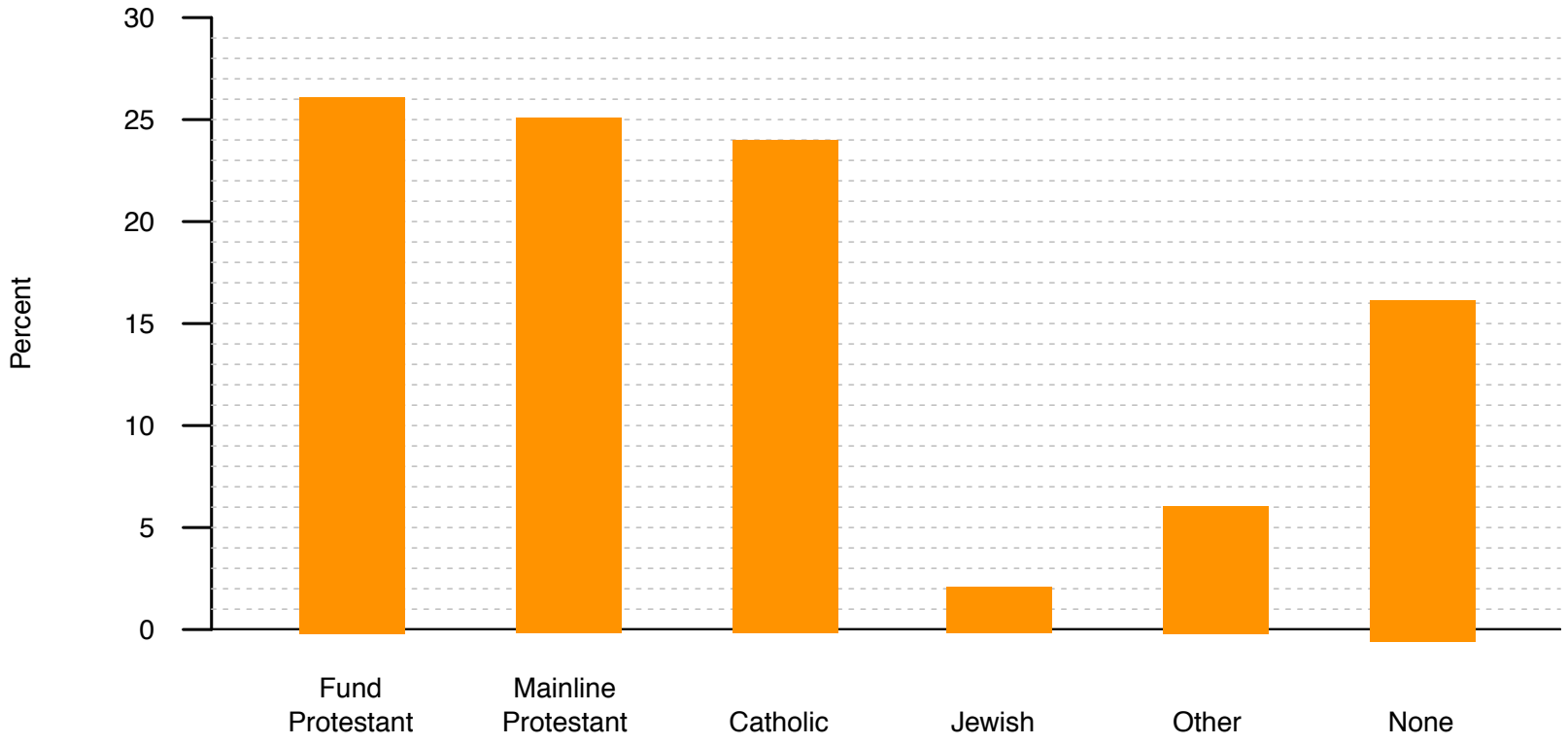
Handout 1: Draw a barplot of the distribution of religious affiliation

```
round(table(sex$relig)/sum(table(sex$relig)),2)*100
```

```
##
##     Fund Protestant Mainline Protestant            Catholic
##                  26                  25                  24
##              Jewish               Other                None
##                   2                   7                  16
```

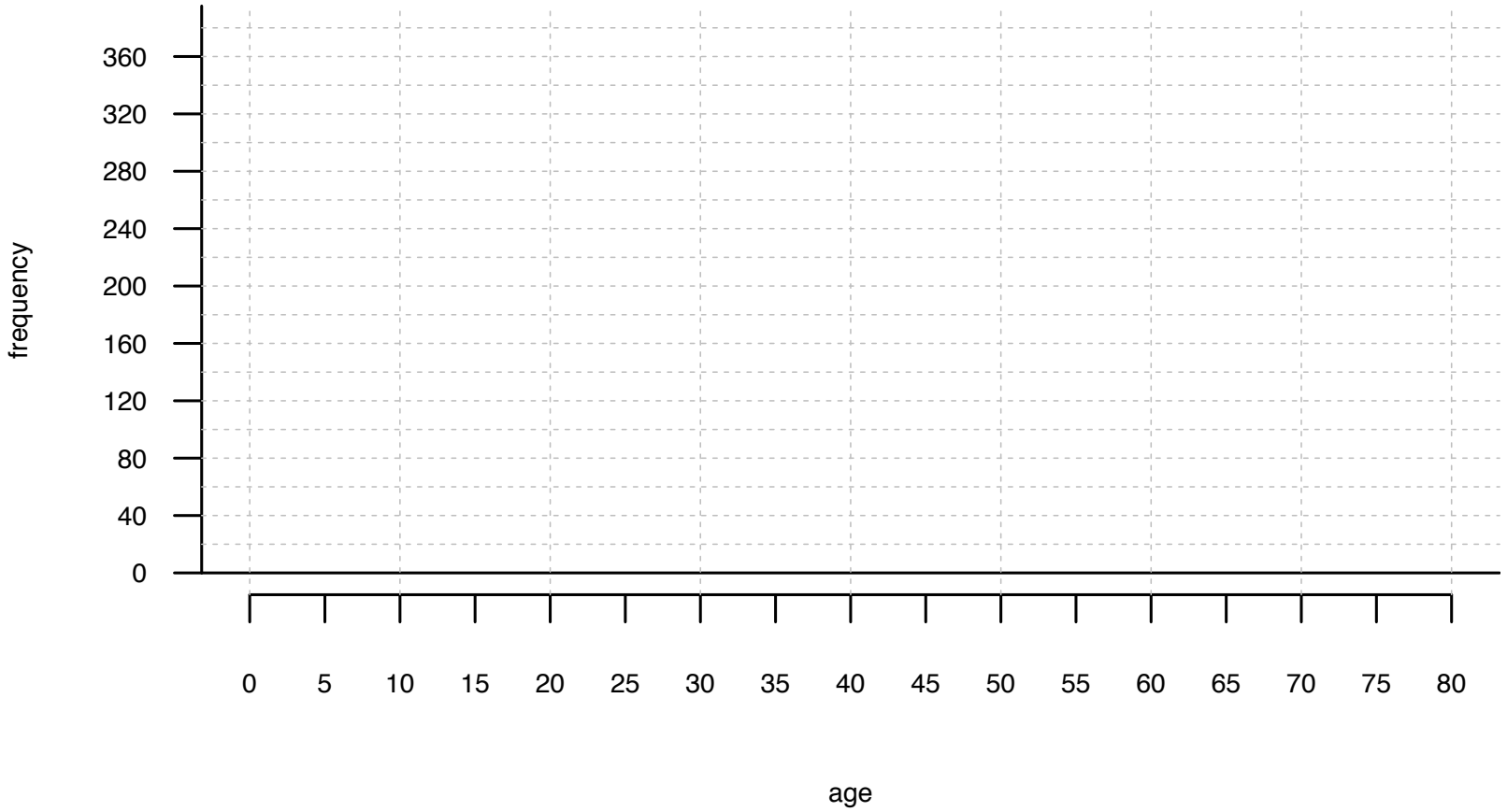


Name (Print and Sign): _____

Handout 2: Draw a histogram of age on the Titanic

```
table(cut(titanic$age, breaks=seq(from=0, to=80, by=10)))
```

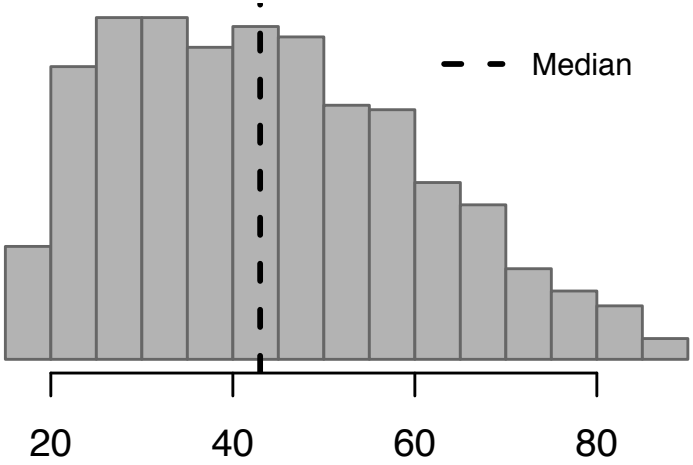
```
##
##  (0,10] (10,20] (20,30] (30,40] (40,50] (50,60] (60,70] (70,80]
##      86     162     361     210     132      62      27       6
```



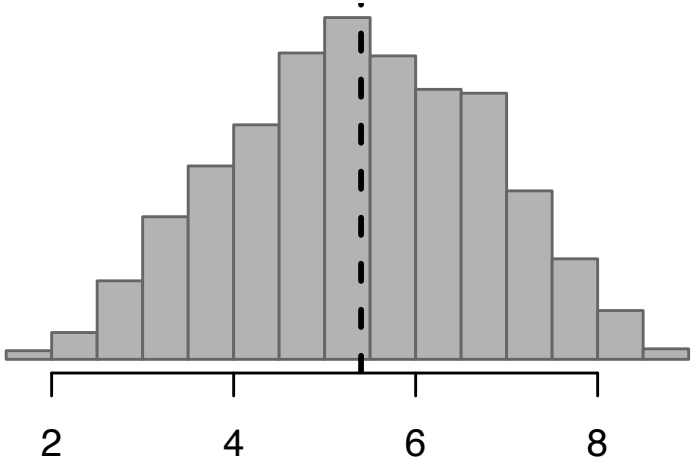
Name (Print and Sign): _____

Handout 3: Estimate and draw the balancing point of each distribution

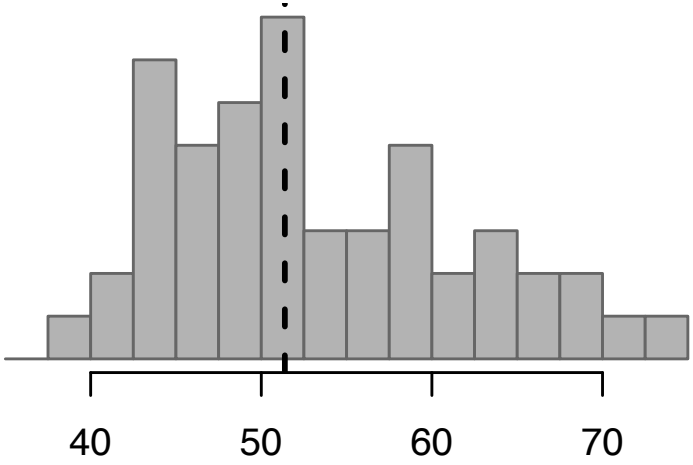
Age of respondents (sex survey)



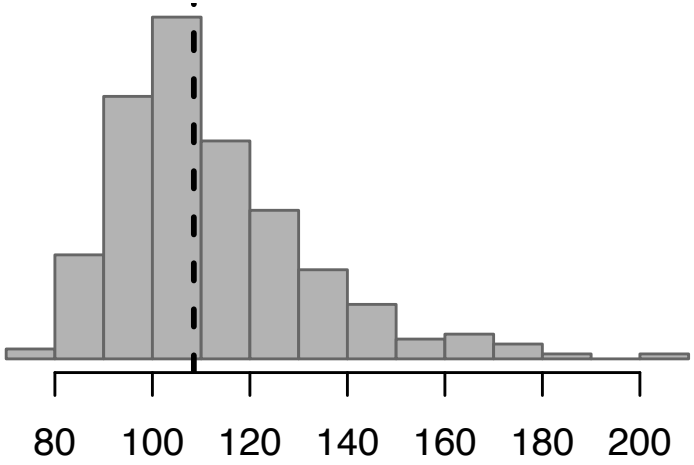
Movie tomato rating



Median income by state (crime data)



SciFi/Fantasy movie runtime

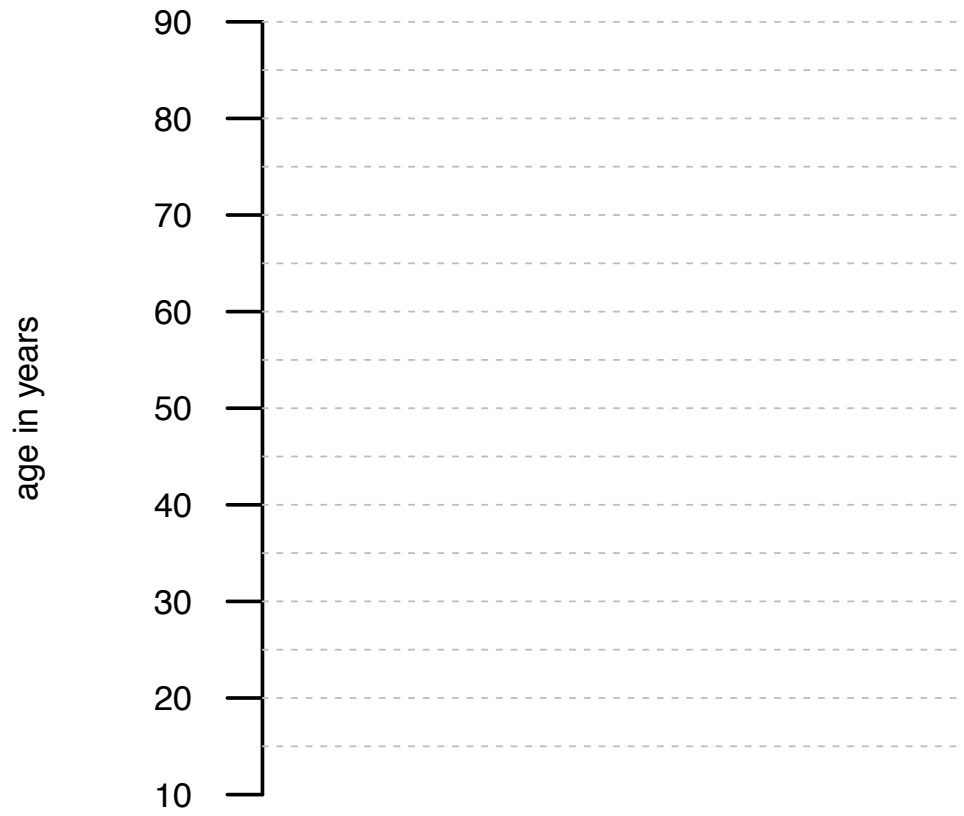


Name (Print and Sign): _____

Handout 4: Draw a boxplot of age distribution

```
quantile(sex$age)
```

```
##   0%  25%  50%  75% 100%
##   18   31   43   56   89
```



Name (Print and Sign): _____

Handout 5: Calculate variance and standard deviation of runtime for 2010 mystery movies

Movie	x	$x - \bar{x}$	$(x - \bar{x})^2$
All Good Things	101		
Edge of Darkness	117		
Wrecked	91		
Sum	309		

$$\bar{x} = 309/3 = 103$$

$$s_x = \sqrt{\sum_{i=1}^{n} (x - \bar{x})^2}/(3 - 1) =$$

Name (Print and Sign): _____

Handout 6: Calculate marginal distributions, conditional distribution of rating by genre, and odds ratio

Rating	**SciFi/Fantasy**	**Action**	**Total**
R or greater	66	106	
PG 13 or less	196	101	
Total			

Distribution of ratings for sciFi/fantasy movies:

Distribution of ratings for action movies:

Odds ratio:

Name (Print and Sign): _____

Handout 6b: Calculate and interpret odds ratio four ways

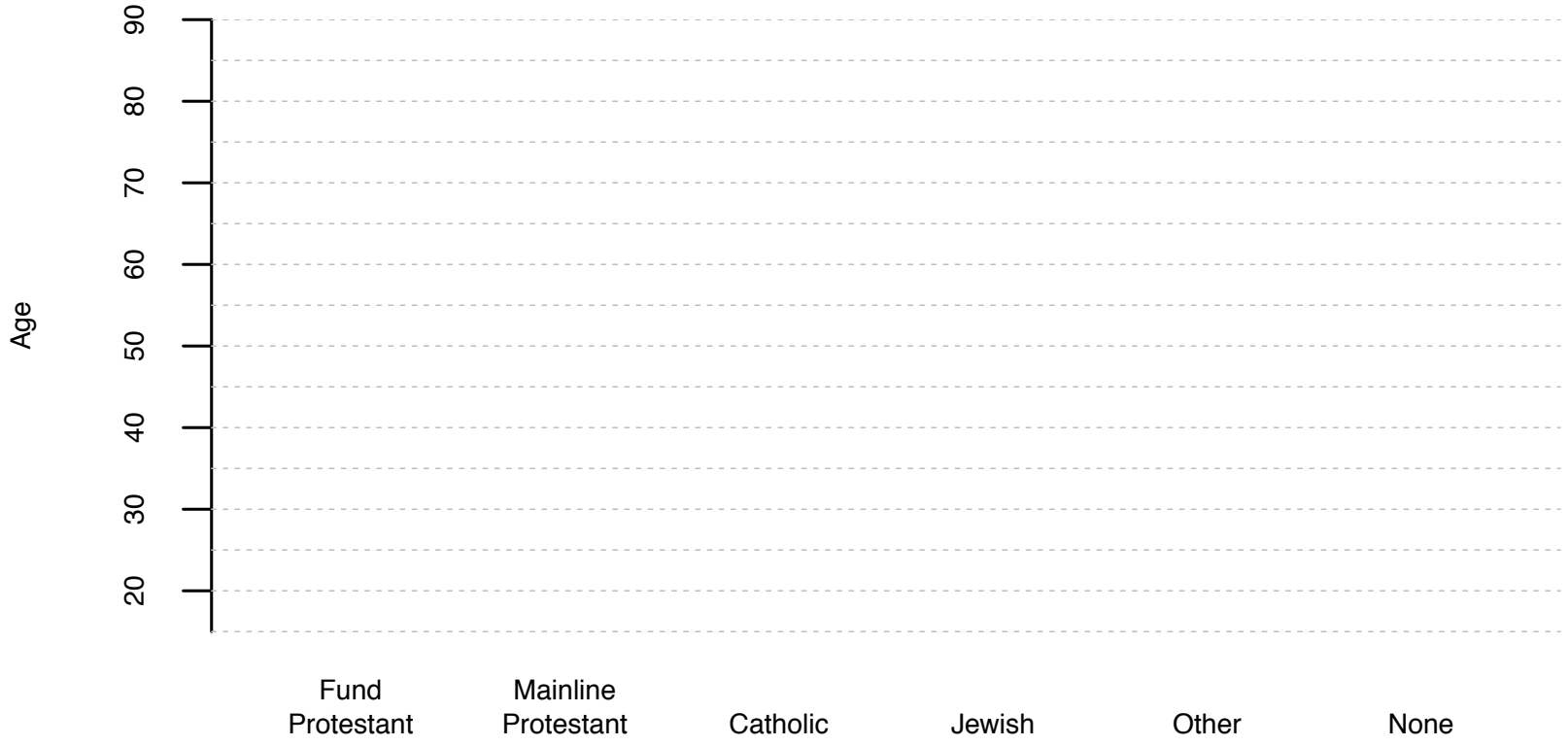
	Odds Ratio	Interpretation

```
##
##                          Internat Domestic
##   Avid soccer fan               3       21
##   Not an avid soccer fan        4       47
##
##                          Domestic Internat
##   Not an avid soccer fan       47        4
##   Avid soccer fan              21        3
##
##                          Domestic Internat
##   Avid soccer fan              21        3
##   Not an avid soccer fan       47        4
##
##                          Internat Domestic
##   Not an avid soccer fan        4       47
##   Avid soccer fan               3       21
```

Name (Print and Sign): _____

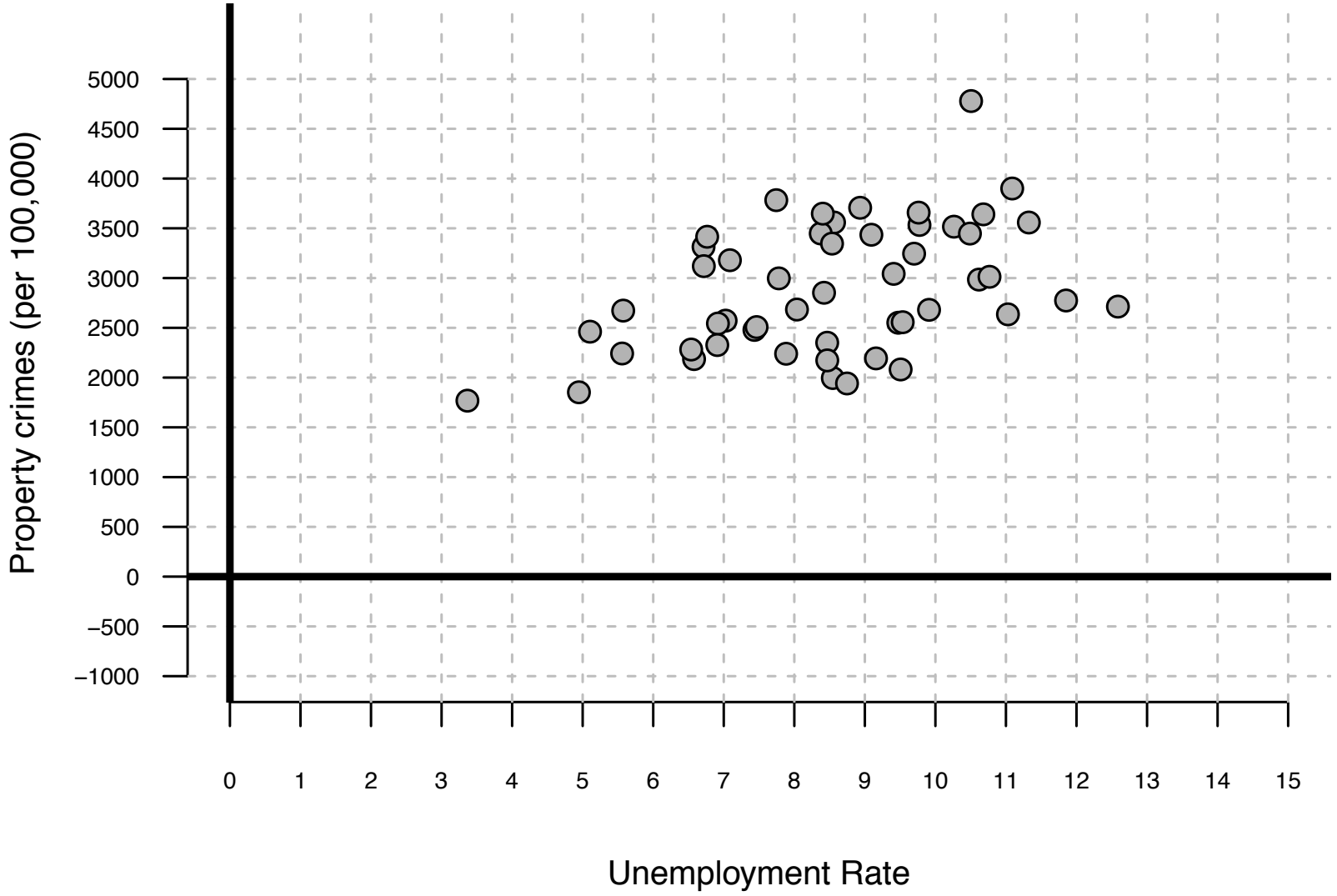
Handout 7: Draw comparative boxplots of age by religious affiliation

	Fund P	Main P	Catholic	Jewish	Other	None
0%	18	18	18	21	18	18
25%	33	32	32	38	28	28
50%	44	46	43	53	37	37
75%	56	60	56	64	48	49
100%	89	89	88	89	77	85



Name (Print and Sign): _____

Handout 8: Draw a straight line through the scatterplot and measure slope and intercept



Name (Print and Sign): _____

Handout 9: Interpret a slope and intercept

```
coef(lm(income~age, data=politics))
```

```
## (Intercept)         age
##  51.3868030   0.1612051
```

Interpret the slope in a single sentence:

Interpret the intercept in a single sentence:

Name (Print and Sign): _____

Handout 10: Confidence interval for proportion supporting gay marriage

```
table(politics$gaymarriage)
```

```
##
## No legal recognition          Civil union Support gay marriage
##                 1515                 1962                 2357
```

```
#sum up to get n
1515+1962+2357
```

```
## [1] 5834
```

```
#estimate p-hat
round(2357/5834,2)
```

```
## [1] 0.4
```

```
#t-stat
qt(0.975, 5834-1)
```

```
## [1] 1.960371
```

$$\texttt{standard error} = \sqrt{\hat{p} * (1 - \hat{p})/n} =$$

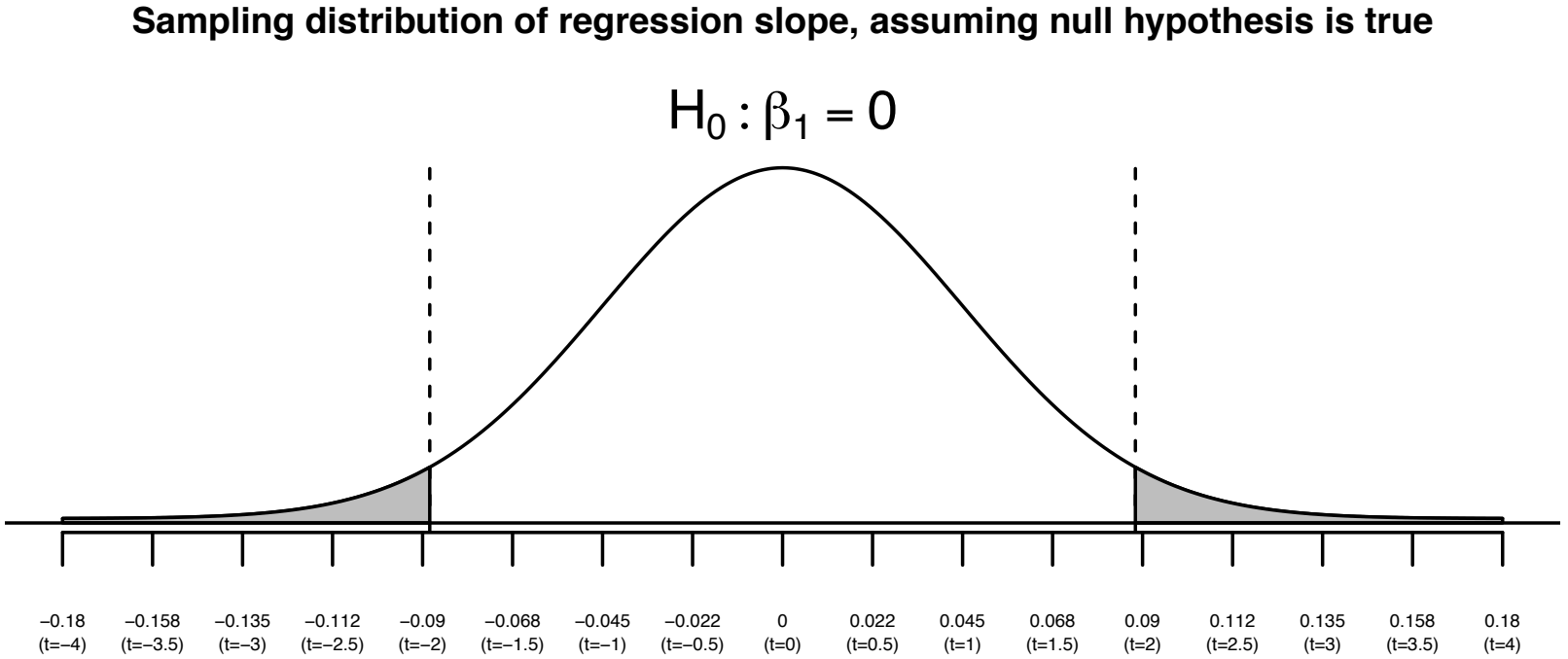
$$\texttt{confidence interval} = \hat{p} \pm t * (\texttt{standard error}) =$$

Name (Print and Sign): _____

Handout 11: Hypothesis test for a slope

```
round(summary(lm(income~age, data=politics))$coef[,1:2],3)
```

```
##             Estimate Std. Error
## (Intercept)   51.387      2.338
## age            0.161      0.045
```



Reject Fail to Reject

Name (Print and Sign): _____

Handout 12: Interpret a slope and intercept from a multivariate regression model

```
round(summary(lm(TomatoMeter~I(Year-2001)+I(Runtime-90)+I(BoxOffice-45), data=movies))$coef,3)
```

```
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         40.891      1.091  37.465    0.000
## I(Year - 2001)       0.422      0.133   3.165    0.002
## I(Runtime - 90)      0.307      0.031   9.826    0.000
## I(BoxOffice - 45)    0.056      0.008   6.983    0.000
```

Interpret the slope on year in a single sentence:

Interpret the slope on runtime in a single sentence:

Interpret the intercept in a single sentence:

Name (Print and Sign): _____

Handout 13: Fill out the "slopes" for the regression models with different references

```
round(tapply(politics$income, politics$race, mean, na.rm=TRUE),0)
```

```
##                   White                   Black                Hispanic
##                      67                      42                      46
## Asian/Pacific Islander        American Indian                   Other
##                      79                      34                      55
```

Variable	Estimate	Estimate
Intercept		
White	(reference)	
Black		(reference)
Hispanic		
Asian		
Am. Indian		

Name (Print and Sign): _____

Handout 14: Interpret the slope and intercept of regression model with categorical predictors

```
round(summary(lm(income~race+educ+I(age-25), data=politics))$coef,2)
```

```
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     34.60       2.63   13.17     0.00
## raceBlack                      -18.32       1.85   -9.92     0.00
## raceHispanic                   -11.93       1.91   -6.23     0.00
## raceAsian/Pacific Islander       0.46       5.01    0.09     0.93
## raceAmerican Indian            -22.92       8.86   -2.59     0.01
## raceOther                      -13.46       7.13   -1.89     0.06
## educHS graduate                 10.32       2.56    4.04     0.00
## educSome college                23.93       2.46    9.73     0.00
## educBA degree                   48.80       2.68   18.19     0.00
## educGrad degree                 67.49       2.92   23.14     0.00
## I(age - 25)                      0.07       0.04    1.72     0.09
```

Interpret the slope on Black in a single sentence:

Interpret the slope on BA degree in a single sentence:

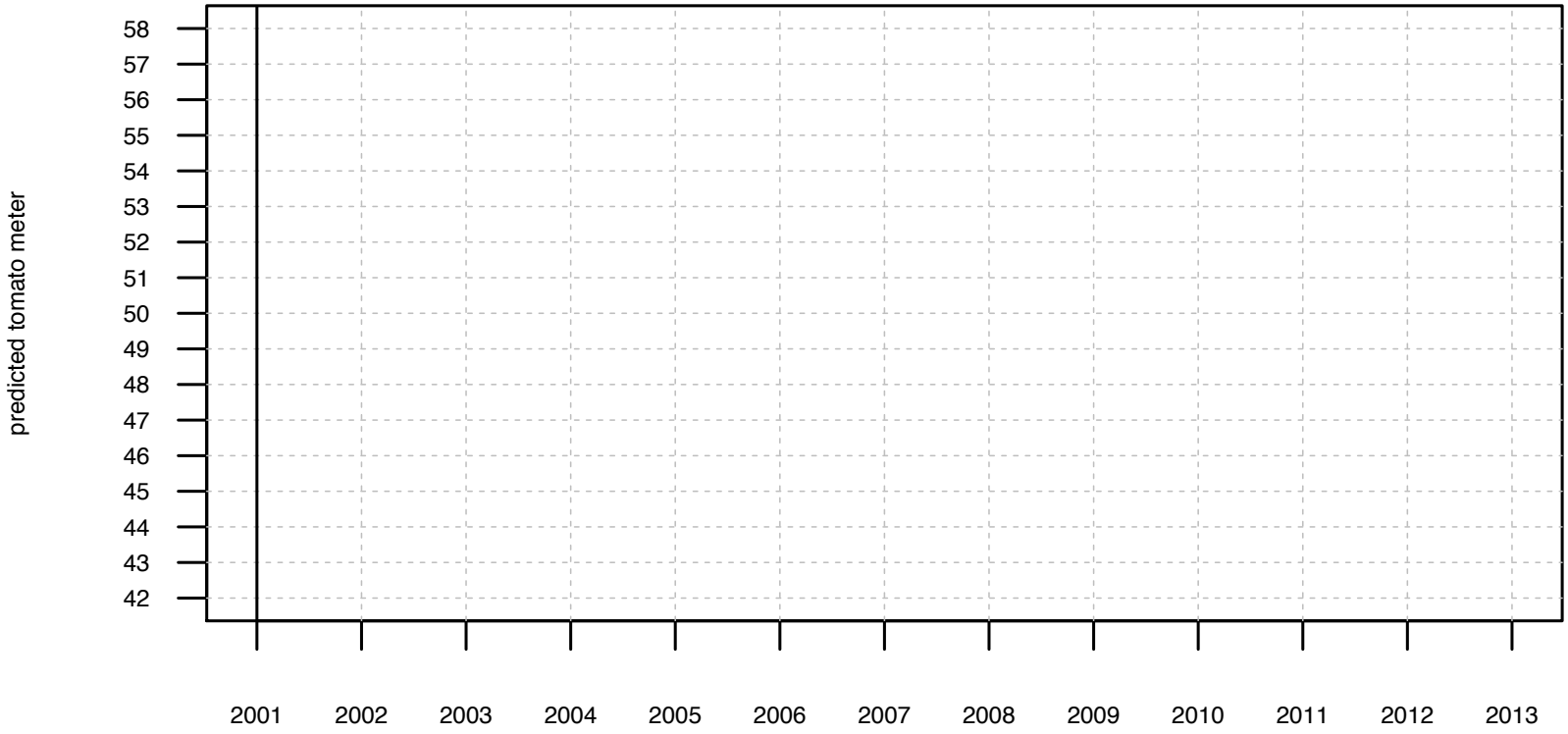
Interpret the intercept in a single sentence:

Name (Print and Sign): _____

Handout 15: Plot two lines showing movie ratings over time for popular and niche movies

```
movies$popular <- movies$Genre=="Action" | movies$Genre=="Animation" |
  movies$Genre=="Comedy" | movies$Genre=="Family" | movies$Genre=="SciFi/Fantasy"
round(summary(lm(movies$TomatoMeter~movies$popular*I(movies$Year-2001)))$coef,2)[,1:2]
```

```
##                                          Estimate Std. Error
## (Intercept)                                 50.58       1.64
## movies$popularTRUE                          -8.00       2.08
## I(movies$Year - 2001)                        0.09       0.22
## movies$popularTRUE:I(movies$Year - 2001)     0.47       0.28
```



Name (Print and Sign): _____

Handout 16: Interpret the slope and intercept of regression model with interactions

```
round(summary(lm(BoxOffice~I(Runtime-90)*Rating, data=movies))$coef,2)
```

```
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       73.99       7.61    9.72     0.00
## I(Runtime - 90)                    2.51       0.52    4.80     0.00
## RatingPG                         -14.64       8.44   -1.74     0.08
## RatingPG-13                      -56.26       8.04   -7.00     0.00
## RatingR                          -61.48       7.96   -7.73     0.00
## RatingNC-17                      -93.21     216.46   -0.43     0.67
## RatingUnrated                    -73.43      10.90   -6.73     0.00
## I(Runtime - 90):RatingPG          -1.65       0.57   -2.91     0.00
## I(Runtime - 90):RatingPG-13       -0.20       0.53   -0.37     0.71
## I(Runtime - 90):RatingR           -1.75       0.53   -3.28     0.00
## I(Runtime - 90):RatingNC-17       -1.20      11.50   -0.10     0.92
## I(Runtime - 90):RatingUnrated     -2.40       0.68   -3.54     0.00
```

Interpret the slope on Runtime in a single sentence:

Interpret the slope on RatingR in a single sentence:

Interpret the interaction term Runtime:RatingR in a single sentence:

Name (Print and Sign): _____