

# Project #1: Grade Calculator

---

## 1. Project Details

### Goal of the project

Calculates the final grade given the weight of each hwk, midterm, final exam and lab report.

The freshmen physics courses at Hunter College of The City University of New York are split into Lecture and Labs. Lecture accounts for 85% of the final grade and labs for 15%. The labs and lecture are taught by different instructors. Combine the data by id number to calculate the final grade.

### The lab grades

- The last two columns represent quiz 1 and quiz 2. Each quiz is graded on the scale of 0-20. Each quiz is worth 10% of the lab grade.
- Columns 1-12 are grades from lab reports. They are graded in the scale of 0-100. The lowest two should be dropped when calculating the final lab grade, take the best 10 out of 12. The lab reports are worth 80% of the lab grade. The students must have minimum of 10 labs to receive a passing grade.

### The lecture grades

- There are two midterms and a final:
  - The lowest midterm is worth 20% and the highest 35%.
  - The final exam is worth 45%.
- 

## 2. Load and clean data

Load necessary libraries, readODS for reading ODS data, xlsx for reading xlsx spreadsheet, dplyr for data wrangling, stringr for string operations.

```
library(readODS)
library(xlsx)
library(dplyr)
library(stringr)
```

First, load the lecture grades from xlsx spread sheet into a data frame named lecture. Change student id from “id” to “ID”, eliminate the while space in the student id.

```
lecture <- read.xlsx2("/Users/qiu/Downloads/DSBootcamp/project1_data/lecture.xlsx",
                     sheetName = "Sheet1",
                     colIndex=c(2:5))
names(lecture) <- names(lecture) %>% str_replace_all("id", "ID") # change from "id" to "ID",
lecture <- mutate(lecture, ID = str_replace_all(ID, "[[ s]]", "")) # remove letter "s" and the white s
lecture
```

```
##      ID MT1 MT2 Final
## 1    1  44  94    50
## 2    2  56  95    75
## 3    3  46  91    43
## 4    4 100  89    62
## 5    5  82  90    42
## 6    6  55  74    43
## 7    7  59  98    64
## 8    8  87  59    64
## 9    9  95  93    74
## 10   10 90  87    78
## 11   11 97  53    57
## 12   12 46  70    52
## 13   13 75  54    52
## 14   14 63  66    64
## 15   15 53  45    74
## 16   16 65  91    85
## 17   17 86  54    81
## 18   18 61  57    76
## 19   19 70  52    69
## 20   20 87  79    53
## 21   21 62  68    69
## 22   22 47  81    42
## 23   23 91  81    70
## 24   24 85  85    50
## 25   25 56  49    55
## 26   26 79  91    56
## 27   27 60  79    43
## 28   28 58  64    47
## 29   29 68  89    80
## 30   30 81  88    50
## 31   31 96  84    68
## 32   32 65  94    82
```

Second, load the lab grades from xlsx spread sheet for lab section 1L12. Notice there is a typo in the name of column “lab5”, which should be “Lab6” instead. Make corrections to the name in the code.

```
lab120sec12 <- read.xlsx("/Users/qiu/Downloads/DSBootcamp/project1_data/Fall 2015 Lab 120, Sec 1L12.xlsx",
  sheetName = "Sheet1",
  rowIndex = c(8:20),
  colIndex = c(1:15))
names(lab120sec12) <- names(lab120sec12) %>%
  str_replace_all("\\\\.", "") %>%
  str_replace("lab5", "LAB6") %>%
  str_replace("StudentsName", "ID")
lab120sec12 <- mutate(lab120sec12, ID = str_replace_all(ID, "s", "")) # remove letter "s" in student i
lab120sec12
```

```
##      ID LAB1 LAB2 LAB3 LAB4 LAB5 LAB6 LAB7 LAB8 LAB9 LAB10 LAB11 LAB12 QUIZ2
## 1   21   70   85   97   90  100   90   80   90   95  100  89.7   NA   6.0
## 2   22   95  100   90   90   90  100   97  100  100  100  96.2   NA  16.0
## 3   23   70   80   90   93   90   90   85   NA   95  100 100.0   NA   6.0
## 4   24   90   70   NA   90   NA   NA   NA   NA   NA   NA   NA   NA   NA
## 5   25   70   90   97   97   95   95   85   NA  100   85 100.0   NA  10.0
```

```
## 6 26 NA NA NA NA NA NA NA NA NA NA NA NA NA
## 7 27 90 85 98 97 90 95 95 100 100 100 NA NA 2.0
## 8 28 90 90 NA 98 100 NA NA NA NA NA NA NA NA
## 9 29 95 85 95 95 95 90 NA 100 90 90 95.0 NA 11.0
## 10 30 88 85 NA NA 97 NA NA NA NA NA NA NA NA
## 11 31 90 90 100 90 90 90 90 98 97 100 100.0 95 6.0
## 12 32 88 100 95 100 100 100 100 100 97 100 100.0 NA 13.5
## QUIZ1
## 1 3
## 2 15
## 3 5
## 4 NA
## 5 10
## 6 NA
## 7 10
## 8 NA
## 9 12
## 10 NA
## 11 6
## 12 7
```

Last, load the lab grades from ods spread sheet for lab section 1L09. The same typo in the name of column “lab5”, which should be “Lab6” instead.

```
ods1 <- read.ods("/Users/qiu/Downloads/DSBootcamp/project1_data/Fall 2015 Lab 120 section 1L09.ods",
                 sheet = 1)
# View(lab120sec09)
lab120sec09 = data.frame(ods1[9:28,1:15])
names(lab120sec09) = c(ods1[8,1:13],ods1[30,14:15])

names(lab120sec09) <- names(lab120sec09) %>%
  str_replace("lab 5","LAB6") %>%
  str_replace("Student's Name","ID") %>%
  str_replace_all("\\\\#", "") %>%
  str_replace_all(" ", "") %>%
  toupper()
lab120sec09 <- mutate(lab120sec09, ID = str_replace_all(ID, "S", ""))
lab120sec09
```

```
## ID LAB1 LAB2 LAB3 LAB4 LAB5 LAB6 LAB7 LAB8 LAB9 LAB10 LAB11 LAB12 QUIZ2
## 1 1 80 90 70
## 2 2 100 100 95 100 100 100 100 100 100 95 100 6
## 3 3 100 100 97 100 100 79 100 100 100 80 97 100 18
## 4 4 80 80 93 90 90 78 95 100 93 90
## 5 5 100 100 93 100 100 86 100 100 95 80 93 100 14
## 6 6 90 95 95 100 80 100 100 95 60 95 95 3
## 7 7 90 85 93 98 80 100 100 100 80 93 98 8
## 8 8 90 90 95 85 100 80 100 95 85 95 85 2
## 9 9 85 95 93 100 100 80 100 100 93 100
## 10 10 90 93 93 90 100 80 95 90 95 93 90 8.5
## 11 11 95 90 93 95 98 80 100 100 100 98 93 95 14.5
## 12 12 100 100 100 100 100 100 100 100 100 100 100 6
## 13 13 85 97 95 98 70 70 100 100 100 95 95 98 12.5
## 14 14 90 80 95 90 70 85 95 90 95 95 95 90 10.5
```

```
## 15 15 70 88 100 100 80 85 95 93 95 95 100 100 14.5
## 16 16 90 80 95 80 85 95 90 95 95 95 95 10
## 17 17 95 90 97 95 80 80 100 100 100 100 97 95 7
## 18 18 85 95 100 60 80 100 100 95 100 100 100 11
## 19 19 90 95 100 80 80 90 95 95 95 95 100 100 13
## 20 20 90 100 93 90 98 100 100 100 93 90 18
## QUIZ1
## 1
## 2 6
## 3 13
## 4 2
## 5 8.5
## 6 5
## 7 1
## 8 5
## 9 6
## 10 2.5
## 11 7
## 12 10
## 13 4
## 14 7.5
## 15 10
## 16 10
## 17 7
## 18 19
## 19 8
## 20 10.5
```

### 3. Process data

Use `rbind` function to join two grades in two separate lab sections vertically, then use `merge` function to joint lecture and lab grades horizontally.

```
lab <- rbind(lab120sec09, lab120sec12)
class_grade <- merge(lecture,lab,by="ID")
class_grade
```

```
## ID MT1 MT2 Final LAB1 LAB2 LAB3 LAB4 LAB5 LAB6 LAB7 LAB8 LAB9 LAB10
## 1 1 44 94 50 80 90 70
## 2 10 90 87 78 90 93 93 90 100 80 95 90 95
## 3 11 97 53 57 95 90 93 95 98 80 100 100 100 98
## 4 12 46 70 52 100 100 100 100 100 100 100 100 100 100
## 5 13 75 54 52 85 97 95 98 70 70 100 100 100 95
## 6 14 63 66 64 90 80 95 90 70 85 95 90 95 95
## 7 15 53 45 74 70 88 100 100 80 85 95 93 95 95
## 8 16 65 91 85 90 80 95 80 85 95 90 95 95
## 9 17 86 54 81 95 90 97 95 80 80 100 100 100 100
## 10 18 61 57 76 85 95 100 60 80 100 100 95 100
## 11 19 70 52 69 90 95 100 80 80 90 95 95 95
## 12 2 56 95 75 100 100 95 100 100 100 100 100 100
## 13 20 87 79 53 90 100 93 90 98 100 100 100
## 14 21 62 68 69 70 85 97 90 100 90 80 90 95 100
## 15 22 47 81 42 95 100 90 90 90 100 97 100 100 100
```

##	16	23	91	81	70	70	80	90	93	90	90	85	<NA>	95	100
##	17	24	85	85	50	90	70	<NA>	90	<NA>	<NA>	<NA>	<NA>	<NA>	<NA>
##	18	25	56	49	55	70	90	97	97	95	95	85	<NA>	100	85
##	19	26	79	91	56	<NA>	<NA>	<NA>	<NA>	<NA>	<NA>	<NA>	<NA>	<NA>	<NA>
##	20	27	60	79	43	90	85	98	97	90	95	95	100	100	100
##	21	28	58	64	47	90	90	<NA>	98	100	<NA>	<NA>	<NA>	<NA>	<NA>
##	22	29	68	89	80	95	85	95	95	95	90	<NA>	100	90	90
##	23	3	46	91	43	100	100	97	100	100	79	100	100	100	80
##	24	30	81	88	50	88	85	<NA>	<NA>	97	<NA>	<NA>	<NA>	<NA>	<NA>
##	25	31	96	84	68	90	90	100	90	90	90	90	98	97	100
##	26	32	65	94	82	88	100	95	100	100	100	100	100	97	100
##	27	4	100	89	62	80	80	93	90	90	78	95	100		
##	28	5	82	90	42	100	100	93	100	100	86	100	100	95	80
##	29	6	55	74	43	90		95	95	100	80	100	100	95	60
##	30	7	59	98	64	90	85	93	98		80	100	100	100	80
##	31	8	87	59	64	90	90	95	85	100	80	100		95	85
##	32	9	95	93	74	85	95	93	100	100	80	100	100		
##		LAB11	LAB12	QUIZ2	QUIZ1										
##	1														
##	2	93	90	8.5	2.5										
##	3	93	95	14.5	7										
##	4	100	100	6	10										
##	5	95	98	12.5	4										
##	6	95	90	10.5	7.5										
##	7	100	100	14.5	10										
##	8	95		10	10										
##	9	97	95	7	7										
##	10	100		11	19										
##	11	100	100	13	8										
##	12	95	100	6	6										
##	13	93	90	18	10.5										
##	14	89.7	<NA>	6	3										
##	15	96.2	<NA>	16	15										
##	16	100	<NA>	6	5										
##	17	<NA>	<NA>	<NA>	<NA>										
##	18	100	<NA>	10	10										
##	19	<NA>	<NA>	<NA>	<NA>										
##	20	<NA>	<NA>	2	10										
##	21	<NA>	<NA>	<NA>	<NA>										
##	22	95	<NA>	11	12										
##	23	97	100	18	13										
##	24	<NA>	<NA>	<NA>	<NA>										
##	25	100	95	6	6										
##	26	100	<NA>	13.5	7										
##	27	93	90		2										
##	28	93	100	14	8.5										
##	29	95	95	3	5										
##	30	93	98	8	1										
##	31	95	85	2	5										
##	32	93	100		6										

#### 4. Analysis data