

Exercises Day 1*

Ronald Geskus

Tran Thai Hung

Last version: April 04, 2025

You might have done exercises 1 and 2 directly on the material slide. For exercises 3 to 7, you will do them in RStudio. Please make sure that you have R and RStudio installed on your laptop.

Table of contents

Exercise 1: selection within a vector	1
Exercise 2: some calculations	2
RStudio	4
Exercise 3: importing data from Excel and first check	5
Exercise 4: some columns summarized	6
Exercise 5: work with subsets	13
Exercise 6: make factor variables	15
Exercise 7: working with dates	16

Exercise 1: selection within a vector

- Create a vector named `vec` that consists of the numbers from 11 to 30
- Select the 7th element of the vector
- Select all elements except the 15th. Hint: use a minus sign
- Select the 2nd and 5th element of the vector
- Select only the odd valued elements of `vec`.

*Please do not circulate.

*Remark: Exercise e. is not an easy exercise. We advise you to do this in two steps. First create a vector **index** that selects the odd numbers in **vec**, using the function **seq**. You may need to consult the help page of **seq** via **help(seq)**.*

```
# a.  
vec <- seq(11,30)  
# b.  
vec[7]
```

```
[1] 17
```

```
# c.  
vec[-15]
```

```
[1] 11 12 13 14 15 16 17 18 19 20 21 22 23 24 26 27 28 29 30
```

```
# d.  
vec[c(2,5)]
```

```
[1] 12 15
```

```
# e.  
index <- seq(1,length(vec),by=2)  
vec[index]
```

```
[1] 11 13 15 17 19 21 23 25 27 29
```

Exercise 2: some calculations

- a. Use the elementary functions $/$, $-$, $\wedge$ and the function **sum** to calculate the mean

$$\bar{x} = \sum_{i=1}^n x_i / n$$

and the standard deviation

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

of the fare paid by the titanic passengers.

b. Verify your answer by using the built-in R functions `mean` and `sd`.

We create the data set that we used in the web based interface.

```
titanic <- data.frame(  
  pclass=rep(c("1st", "2nd", "3rd"), c(4, 2, 4)),  
  survived=c(1, 0, 0, 0, 1, 1, 0, 0, 0, 0),  
  name=c("Allison, Master. Hudson Trevor", "Dulles, Mr. William Crothers", "McCarthy,  
  age= c(0.9167, 39, 54, 47, 27, 13, 13, 17, NA, NA),  
  fare=c(151.55, 29.7, 51.8625, 34.0208, 13.8583, 19.5, 20.25, 8.6625, 7.75, 23.4))
```

```
# a.  
fare <- titanic$fare  
MeanFare <- sum(fare)/length(fare)  
MeanFare
```

```
[1] 36.05541
```

```
# standard deviation (in three baby steps)  
numerator <- sum((fare-MeanFare)^2)  
denominator <- (length(fare)-1)  
StdFare <- sqrt(numerator/denominator)  
StdFare
```

```
[1] 42.63766
```

```
# b.  
mean(fare)
```

```
[1] 36.05541
```

```
sd(fare)
```

```
[1] 42.63766
```

RStudio

We will work with the dataset that gives the survival status of passengers on the Titanic. We provide the dataset in **Excel** format ([Titanic3.xlsx](#)). See [this link](#) for a description of the variables. Our data set has three additional columns: **dob** (date of birth), **family** (whether the person travelled with family) and **agecat** (age categorized in age ranges).

We assume that you have created a folder on your computer for the R course. Start the RStudio program from within that folder. Your working directory should be that folder. If it is not, set the working directory via

Session → Set Working Directory → Choose Directory...

Open a new R Script file via the menu:

File → New File → R Script...

Write the code in that R Script file.

Some useful Keyboard Shortcuts in RStudio for Windows

Over time it is useful to learn a couple of RStudio shortcuts that you frequently use. You can find an overview of them under the **Help → Keyboard Shortcuts help** menu. More shortcuts and tips can be found on the [Appsilon](#) website.

Note

On the mac, Ctrl → Cmd, Alt → Option, and Enter → Return.

- **Ctrl+Enter**: Run current line
- **Ctrl+Alt+I**: Insert new chunk
- **Alt+-**: Insert assignment <-
- **F1**: Show function help page when cursor is on function name
- **Ctrl+Tab**: Switch between tabs in Source Editor
- **Ctrl+1 / Ctrl+2**: Move focus to Source/Console
- **Tab / Ctrl+Space**: Code suggestion/completion
- **In Console**:
 - Arrows **Up/Down** to navigate earlier commands
 - **Ctrl+Up** pops up a listing of latest commands

Exercise 3: importing data from Excel and first check

Import the dataset by clicking on **Import Dataset** from the **Environment** pane. Browse for the Excel file. Assign the R object name `titanic` to the dataset under **Import Options**. Copy the code under **Code Preview**. Click on the “Import” button. Paste the code you have copied earlier to your **Script File**.

- Inspect the data set in spreadsheet-like format that shows up after importing (if it does not show up, you can click on the `titanic` object in the **Environment** pane). Do the variables make sense? Are there any weird values? Sort the data in ascending order of age by clicking on the **age** in the data set.
- Use the function `dim` to find the number of rows (passengers) and columns (variables) in the data set.

```
dim(titanic)
```

```
[1] 1309  17
```

Answer: 1309 rows and 17 columns

- Use the `str` function to obtain a summary of the basic characteristics per variable. Have a look at the columns for `survived` and `age`. Do they have the mode that you expect? (You can also use the `mode` function to check their modes.)

```
str(titanic)
```

```
tibble [1,309 x 17] (S3: tbl_df/tbl/data.frame)
 $ pclass   : chr [1:1309] "1st" "1st" "1st" "1st" ...
 $ survived : num [1:1309] 1 1 0 0 0 1 1 0 1 0 ...
 $ name     : chr [1:1309] "Allen, Miss. Elisabeth Walton" "Allison, Master. Hudson Tr
 $ sex      : chr [1:1309] "female" "male" "female" "male" ...
 $ age      : chr [1:1309] "29" "0.91670000600000001" "2" "30" ...
 $ sibsp    : num [1:1309] 0 1 1 1 1 0 1 0 2 0 ...
 $ parch    : num [1:1309] 0 2 2 2 2 0 0 0 0 0 ...
 $ ticket   : chr [1:1309] "24160" "113781" "113781" "113781" ...
 $ fare     : num [1:1309] 211 152 152 152 152 ...
 $ cabin    : chr [1:1309] "B5" "C22 C26" "C22 C26" "C22 C26" ...
 $ embarked : chr [1:1309] "Southampton" "Southampton" "Southampton" "Southampton" ...
 $ boat     : chr [1:1309] "2" "11" NA NA ...
```

```
$ body      : chr [1:1309] "NA" "NA" "NA" "135" ...
$ home_dest: chr [1:1309] "St Louis, MO" "Montreal, PQ / Chesterville, ON" "Montreal,
$ dob       : chr [1:1309] "-28019.25" "-17761.8246771915" "-18157.5" "-28384.5" ...
$ family    : chr [1:1309] "no" "yes" "yes" "yes" ...
$ agecat    : chr [1:1309] "(18,30]" "(0,18]" "(0,18]" "(18,30]" ...
```

```
mode(titanic$survived)
```

```
[1] "numeric"
```

```
mode(titanic$age)
```

```
[1] "character"
```

`survived` is of mode numeric, while one would expect it to be a categorical variable of mode character. `age` is of mode character, while one would expect it to be of mode numeric

Exercise 4: some columns summarized

A problem with the mode for age is that missing values are coded as a character value "NA" in our Excel file; it is not recognized as an R type NA value. In the future, you may encounter other representations of missing values, such as `missing`, `404`, `-99` or `unknown`, which should also be treated as NA. Read in the data set as `titanic` object again, but specify NA as the missing value indicator under **Code Preview**. Your code under **Code Preview** should now look like

```
library(readxl)
titanic <- read_excel("data/raw/Titanic3.xlsx", na = "NA")
View(titanic)
```

- a. Issue the commands

```
head(titanic, 10)
```

```
# A tibble: 10 x 17
  pclass survived name      sex      age sibsp parch ticket  fare cabin embarked
  <chr>      <dbl> <chr>      <chr>  <dbl> <dbl> <dbl> <chr>  <dbl> <chr> <chr>
1 1st          1 Allen, ~ fema~ 29          0      0 24160  211. B5      Southam~
2 1st          1 Allison~ male   0.917      1      2 113781 152. C22 ~ Southam~
3 1st          0 Allison~ fema~ 2          1      2 113781 152. C22 ~ Southam~
4 1st          0 Allison~ male   30          1      2 113781 152. C22 ~ Southam~
5 1st          0 Allison~ fema~ 25          1      2 113781 152. C22 ~ Southam~
6 1st          1 Anderso~ male   48          0      0 19952   26.5 E12      Southam~
7 1st          1 Andrews~ fema~ 63          1      0 13502   78.0 D7       Southam~
8 1st          0 Andrews~ male   39          0      0 112050   0 A36      Southam~
9 1st          1 Appleto~ fema~ 53          2      0 11769   51.5 C101     Southam~
10 1st          0 Artagav~ male   71          0      0 PC 17~  49.5 <NA>   Cherbou~
# i 6 more variables: boat <chr>, body <dbl>, home_dest <chr>, dob <dbl>,
#   family <chr>, agecat <chr>
```

```
tail(titanic, 10)
```

```
# A tibble: 10 x 17
  pclass survived name      sex      age sibsp parch ticket  fare cabin embarked
  <chr>      <dbl> <chr>      <chr>  <dbl> <dbl> <dbl> <chr>  <dbl> <chr> <chr>
1 3rd          0 Yasbeck,~ male   27          1      0 2659   14.5 <NA>   Cherbou~
2 3rd          1 Yasbeck,~ fema~ 15          1      0 2659   14.5 <NA>   Cherbou~
3 3rd          0 Yousseff,~ male   45.5      0      0 2628    7.22 <NA>   Cherbou~
4 3rd          0 Yousif, ~ male   NA          0      0 2647    7.22 <NA>   Cherbou~
5 3rd          0 Yousseff~ male   NA          0      0 2627   14.5 <NA>   Cherbou~
6 3rd          0 Zabour, ~ fema~ 14.5      1      0 2665   14.5 <NA>   Cherbou~
7 3rd          0 Zabour, ~ fema~ NA          1      0 2665   14.5 <NA>   Cherbou~
8 3rd          0 Zakarian~ male   26.5      0      0 2656    7.22 <NA>   Cherbou~
9 3rd          0 Zakarian~ male   27          0      0 2670    7.22 <NA>   Cherbou~
10 3rd          0 Zimmerma~ male   29          0      0 315082  7.88 <NA>   Southam~
# i 6 more variables: boat <chr>, body <dbl>, home_dest <chr>, dob <dbl>,
#   family <chr>, agecat <chr>
```

What do you observe?

Answer: First and last to rows are shown. Not all columns are shown.

- b. The `readxl` package that is used to import the data stores the data as a `tibble`, which is a special type of format on top of the `data.frame` format to store data in R. Override the `tibble` format using the function `as.data.frame`.

```
titanic <- as.data.frame(titanic)
```

Run the head and tail functions again. What has changed?

```
head(titanic, 10)
```

	pclass	survived	name	sex	age	sibsp	parch
1	1st	1	Allen, Miss. Elisabeth Walton	female	29.0000	0	0
2	1st	1	Allison, Master. Hudson Trevor	male	0.9167	1	2
3	1st	0	Allison, Miss. Helen Loraine	female	2.0000	1	2
4	1st	0	Allison, Mr. Hudson Joshua Crei	male	30.0000	1	2
5	1st	0	Allison, Mrs. Hudson J C (Bessi	female	25.0000	1	2
6	1st	1	Anderson, Mr. Harry	male	48.0000	0	0
7	1st	1	Andrews, Miss. Kornelia Theodos	female	63.0000	1	0
8	1st	0	Andrews, Mr. Thomas Jr	male	39.0000	0	0
9	1st	1	Appleton, Mrs. Edward Dale (Cha	female	53.0000	2	0
10	1st	0	Artagaveytia, Mr. Ramon	male	71.0000	0	0

	ticket	fare	cabin	embarked	boat	body
1	24160	211.3375	B5	Southampton	2	NA
2	113781	151.5500	C22 C26	Southampton	11	NA
3	113781	151.5500	C22 C26	Southampton	<NA>	NA
4	113781	151.5500	C22 C26	Southampton	<NA>	135
5	113781	151.5500	C22 C26	Southampton	<NA>	NA
6	19952	26.5500	E12	Southampton	3	NA
7	13502	77.9583	D7	Southampton	10	NA
8	112050	0.0000	A36	Southampton	<NA>	NA
9	11769	51.4792	C101	Southampton	D	NA
10	PC 17609	49.5042	<NA>	Cherbourg	<NA>	22

	home_dest	dob	family	agecat
1	St Louis, MO	-28019.25	no	(18,30]
2	Montreal, PQ / Chesterville, ON	-17761.82	yes	(0,18]
3	Montreal, PQ / Chesterville, ON	-18157.50	yes	(0,18]
4	Montreal, PQ / Chesterville, ON	-28384.50	yes	(18,30]
5	Montreal, PQ / Chesterville, ON	-26558.25	yes	(18,30]
6	New York, NY	-34959.00	no	(40,80]
7	Hudson, NY	-40437.75	yes	(40,80]
8	Belfast, NI	-31671.75	no	(30,40]
9	Bayside, Queens, NY	-36785.25	yes	(40,80]
10	Montevideo, Uruguay	-43359.75	no	(40,80]


```
tail(titanic, 10)
```

	pclass	survived		name	sex	age	sibsp	parch
1300	3rd	0		Yasbeck, Mr. Antoni	male	27.0	1	0
1301	3rd	1	Yasbeck, Mrs. Antoni (Selini Al	female	15.0		1	0
1302	3rd	0		Youseff, Mr. Gerious	male	45.5	0	0
1303	3rd	0		Yousif, Mr. Wazli	male	NA	0	0
1304	3rd	0		Yousseff, Mr. Gerious	male	NA	0	0
1305	3rd	0		Zabour, Miss. Hileni	female	14.5	1	0
1306	3rd	0		Zabour, Miss. Thamine	female	NA	1	0
1307	3rd	0	Zakarian, Mr. Mapriededer	male	26.5		0	0
1308	3rd	0		Zakarian, Mr. Ortin	male	27.0	0	0
1309	3rd	0		Zimmerman, Mr. Leo	male	29.0	0	0

	ticket	fare	cabin	embarked	boat	body	home_dest	dob	family
1300	2659	14.4542	<NA>	Cherbourg	C	NA	<NA>	-27288.75	yes
1301	2659	14.4542	<NA>	Cherbourg	<NA>	NA	<NA>	-22905.75	yes
1302	2628	7.2250	<NA>	Cherbourg	<NA>	312	<NA>	-34045.88	no
1303	2647	7.2250	<NA>	Cherbourg	<NA>	NA	<NA>	NA	no
1304	2627	14.4583	<NA>	Cherbourg	<NA>	NA	<NA>	NA	no
1305	2665	14.4542	<NA>	Cherbourg	<NA>	328	<NA>	-22723.12	yes
1306	2665	14.4542	<NA>	Cherbourg	<NA>	NA	<NA>	NA	yes
1307	2656	7.2250	<NA>	Cherbourg	<NA>	304	<NA>	-27106.12	no
1308	2670	7.2250	<NA>	Cherbourg	<NA>	NA	<NA>	-27288.75	no
1309	315082	7.8750	<NA>	Southampton	<NA>	NA	<NA>	-28019.25	no

	agecat
1300	(18,30]
1301	(0,18]
1302	(40,80]
1303	<NA>
1304	<NA>
1305	(0,18]
1306	<NA>
1307	(18,30]
1308	(18,30]
1309	(18,30]

Answer: All columns are shown.

- c. Summarize the whole dataset using the `summary` function. What do you observe for each of the variables?

```
summary(titanic)
```

pclass	survived	name	sex
Length:1309	Min. :0.000	Length:1309	Length:1309
Class :character	1st Qu.:0.000	Class :character	Class :character
Mode :character	Median :0.000	Mode :character	Mode :character
	Mean :0.382		
	3rd Qu.:1.000		
	Max. :1.000		

age	sibsp	parch	ticket
Min. : 0.1667	Min. :0.0000	Min. :0.000	Length:1309
1st Qu.:21.0000	1st Qu.:0.0000	1st Qu.:0.000	Class :character
Median :28.0000	Median :0.0000	Median :0.000	Mode :character
Mean :29.8811	Mean :0.4989	Mean :0.385	
3rd Qu.:39.0000	3rd Qu.:1.0000	3rd Qu.:0.000	
Max. :80.0000	Max. :8.0000	Max. :9.000	
NA's :263			

fare	cabin	embarked	boat
Min. : 0.000	Length:1309	Length:1309	Length:1309
1st Qu.: 7.896	Class :character	Class :character	Class :character
Median :14.454	Mode :character	Mode :character	Mode :character
Mean : 33.295			
3rd Qu.:31.275			
Max. :512.329			
NA's :1			

body	home_dest	dob	family
Min. : 1.0	Length:1309	Min. : -46647	Length:1309
1st Qu.:72.0	Class :character	1st Qu.: -31672	Class :character
Median :155.0	Mode :character	Median : -27654	Mode :character
Mean :160.8		Mean : -28341	
3rd Qu.:256.0		3rd Qu.: -25097	
Max. :328.0		Max. : -17488	
NA's :1188		NA's :263	

agecat
Length:1309
Class :character
Mode :character

Many are of class and mode character, and the `summary` function does not give any information. The `dob` variable has negative values.

- d. Compute the 5%, 25%, 50%, 75% and 95% quantiles, the inter-quartile range and the standard deviation for age and fare. You may have to take a look at the `help` files for the functions `quantile`, `IQR` and `sd`.

```
## some quantiles
quantile(titanic$age, probs=c(0.05,0.25,0.5,0.75,0.95), na.rm=TRUE)
```

```
5% 25% 50% 75% 95%
5  21  28  39  57
```

```
quantile(titanic$fare, probs=c(0.05,0.25,0.5,0.75,0.95), na.rm=TRUE)
```

```
5%      25%      50%      75%      95%
7.2250  7.8958  14.4542  31.2750 133.6500
```

```
## the IQR (note that this is a single number)
IQR(titanic$age, na.rm=TRUE)
```

```
[1] 18
```

```
IQR(titanic$fare, na.rm=TRUE)
```

```
[1] 23.3792
```

```
## the the standard deviation
sd(titanic$age, na.rm = TRUE)
```

```
[1] 14.4135
```

```
sd(titanic$fare, na.rm = TRUE)
```

```
[1] 51.75867
```

- e. For categorical variables, `summary` is not very informative. Use the `table` function to summarize the variables `pclass`, and `sex`. Also give a two-by-two table for sex and survival status using the same `table` function. Again, have a look at the help file if needed.

```
## basic tables  
table(titanic$pclass)
```

```
1st 2nd 3rd  
323 277 709
```

```
table(titanic$sex)
```

```
female  male  
   466    843
```

```
table(titanic$sex, titanic$survived)
```

```
      0  1  
female 127 339  
male   682 161
```

```
# has no added value because no missings  
table(titanic$sex, titanic$survived, useNA = "always")
```

```
      0  1 <NA>  
female 127 339  0  
male   682 161  0  
<NA>    0  0  0
```

- f. The function `addmargins` adds the row sums and the column sums to the table of survival status by sex. The function `proportions` tabulates proportions instead of absolute numbers.

```
addmargins(table(titanic$sex, titanic$survived))
```

	0	1	Sum
female	127	339	466
male	682	161	843
Sum	809	500	1309

```
proportions(table(titanic$sex, titanic$survived))
```

	0	1
female	0.09702063	0.25897632
male	0.52100840	0.12299465

Use the `proportions` function to compute the fraction (or percentage) of survivors by sex.

```
## basic tables
proportions(table(titanic$sex, titanic$survived), margin = 1)
```

	0	1
female	0.2725322	0.7274678
male	0.8090154	0.1909846

Exercise 5: work with subsets

- Take a look at the name, and the home town/destination of all passengers who were older than 70 years. Use the appropriate selection functions.

```
subset(titanic, age > 70)[, c("name", "home_dest")]
```

	name	home_dest
10	Artagaveytia, Mr. Ramon	Montevideo, Uruguay
15	Barkworth, Mr. Algernon Henry W	Hessle, Yorks
62	Cavendish, Mrs. Tyrell William Little Onn	Hall, Staffs
136	Goldschmidt, Mr. George B	New York, NY
728	Connors, Mr. Patrick	<NA>
1236	Svensson, Mr. Johan	<NA>

It is seen that there is one person from Uruguay in this group. Select the single record from that person. Did this person travel with relatives?

```
subset(titanic, name=="Artagaveytia, Mr. Ramon")
```

```

      pclass survived      name sex age sibsp parch ticket
10      1st         0 Artagaveytia, Mr. Ramon male  71     0     0 PC 17609
      fare cabin embarked boat body      home_dest      dob family
10 49.5042  <NA> Cherbourg <NA>   22 Montevideo, Uruguay -43359.75     no
      agecat
10 (40,80]
```

Answer: the person did not travel with relatives

- b. Make a table of survivor status by sex, but only for the first class passengers. What do you observe? Compare this with the survival status for the 3rd class passengers.

```
first <- xtabs(~sex+survived, data=titanic, subset=(pclass=="1st"))
first
```

```

      survived
sex      0    1
female   5 139
male    118  61
```

```
third <- xtabs(~sex+survived, data=titanic, subset=(pclass=="3rd"))
third
```

```

      survived
sex      0    1
female 110 106
male   418  75
```

Answer: females were much more likely to survive than males, especially among those who travelled first class

Exercise 6: make factor variables

- Make passenger class and sex into factor variables.
- Add a variable `status` to the Titanic data set, which gives the survival status of the passengers as a factor variable with labels “no” and “yes” describing whether individuals survived. Make the value “no” the first level.

```
titanic$status <- factor(titanic$survived, labels=c("no","yes"))
```

- Run the `summary` function again. What change do you observe compared to Exercise 4?

```
summary(titanic)
```

pclass	survived	name	sex
Length:1309	Min. :0.000	Length:1309	Length:1309
Class :character	1st Qu.:0.000	Class :character	Class :character
Mode :character	Median :0.000	Mode :character	Mode :character
	Mean :0.382		
	3rd Qu.:1.000		
	Max. :1.000		

age	sibsp	parch	ticket
Min. : 0.1667	Min. :0.0000	Min. :0.000	Length:1309
1st Qu.:21.0000	1st Qu.:0.0000	1st Qu.:0.000	Class :character
Median :28.0000	Median :0.0000	Median :0.000	Mode :character
Mean :29.8811	Mean :0.4989	Mean :0.385	
3rd Qu.:39.0000	3rd Qu.:1.0000	3rd Qu.:0.000	
Max. :80.0000	Max. :8.0000	Max. :9.000	
NA's :263			

fare	cabin	embarked	boat
Min. : 0.000	Length:1309	Length:1309	Length:1309
1st Qu.: 7.896	Class :character	Class :character	Class :character
Median :14.454	Mode :character	Mode :character	Mode :character
Mean : 33.295			
3rd Qu.:31.275			
Max. :512.329			
NA's :1			

body	home_dest	dob	family
------	-----------	-----	--------

Min. : 1.0	Length:1309	Min. : -46647	Length:1309
1st Qu.: 72.0	Class :character	1st Qu.: -31672	Class :character
Median :155.0	Mode :character	Median : -27654	Mode :character
Mean :160.8		Mean : -28341	
3rd Qu.:256.0		3rd Qu.: -25097	
Max. :328.0		Max. : -17488	
NA's :1188		NA's :263	

agecat	status
Length:1309	no :809
Class :character	yes:500
Mode :character	

Answer: Passenger class and sex are summarized properly, and status is summarized as a categorical variable.

Exercise 7: working with dates

- The variable `dob` gives the date of birth, using days since January 1st, 1960 (this is the time origin used by Stata). Transform this to a true date value using the function `as.Date` (specify the origin argument).

```
titanic$dob <- as.Date(titanic$dob, origin = "1960/1/1")
```

- The default display is “year/month/day”. Make a table of the day of the month that the passengers were born, using the `format` function. What do you observe?

```
table(format(titanic$dob, "%d"))
```

```
13 14 15 16
15 940 86 5
```

Answer: All have a day in the middle of the month

- c. What was the earliest date of birth among all the passengers on the boat. Does this correspond to the age of the oldest person on the date that the ship sank (April 15th, 1912)? Use the functions `min` and `max`.

```
min(titanic$dob, na.rm = TRUE)
```

```
[1] "1832-04-14"
```

```
max(titanic$age, na.rm = TRUE)
```

```
[1] 80
```

```
min(titanic$dob, na.rm = TRUE) + max(titanic$age, na.rm = TRUE) * 365.25
```

```
[1] "1912-04-15"
```

Yes, it does correspond