

MATH 18/19: Analysis of ILA assessments

George Kinnear

28/11/2020

Read data

Read in the `data_math-1819.csv` file which contains the codes assigned to all exam questions, and filter to just the ILA ones.

Sample of the data:

Component	Week	MATH_Group	total_marks_for_Group	total_marks_for_paper	prop_marks
Exam	2011Dec	C	18	120	0.1500000
Exam	2012Aug	C	20	120	0.1666667
Exam	2012Dec	C	18	120	0.1500000
Exam	2013Aug	C	35	120	0.2916667
Exam	2013Dec	C	24	120	0.2000000
Exam	2014Aug	C	12	120	0.1000000

Add the data on coursework, where the ratings were all done in the IRR data file:

```
cw_ratings = read.csv(here::here("data_irr.csv"), header = TRUE, stringsAsFactors = FALSE) %>%
  filter(
    Course == "ILA",
    !str_detect(Assessment, "Exam")
  ) %>%
  select(Assessment, Question, Marks, Agreed) %>%
  mutate(
    Component = str_sub(Assessment,1,1),
    Component = case_when(
      Component == "O" ~ "Online",
      Component == "W" ~ "Written",
      Component == "R" ~ "Reading"
    ),
    Week = parse_number(Assessment),
    Item = paste(Assessment, Question),
    MATH = fct_relevel(fct_expand(as.factor(str_trim(Agreed))), mathcats), mathcats),
    MATH_Group = as.factor(str_sub(MATH,1,1))
  ) %>%
  mutate(
    Marks = replace_na(Marks, 1)
  ) %>%
  # get rid of questions which were coded but had been dropped from Written Assignemnts
  filter( Marks > 0 ) %>%
  select(Component, Week, Item, Marks, MATH, MATH_Group)

cw_ratings %>%
  head() %>%
  kable(booktabs = T)
```

Component	Week	Item	Marks	MATH	MATH_Group
Online	1	Online1 1	1	B1	B
Online	1	Online1 2	1	A2	A
Online	1	Online1 3	1	A3	A
Online	1	Online1 4	1	A3	A
Online	1	Online1 5	1	C1	C
Online	2	Online2 1	1	A3	A

```
cw_item_totals = cw_ratings %>%
  group_by(Component,Week) %>%
  summarise(
    total_marks_for_paper = sum(Marks)
  )
```

```
## `summarise()` regrouping output by 'Component' (override with `.groups` argument)
```

```
cw_item_totals %>%
  pivot_wider(
    names_from = Component,
    values_from = total_marks_for_paper
  ) %>%
  kable(booktabs = TRUE, caption = "Total marks in each assessment")
```

Total marks in each assessment

Week	Online	Reading	Written
1	5	4	10
2	7	4	10

Week	Online	Reading	Written
3	4	4	10
4	5	5	10
5	6	6	10
6	6	4	10
7	4	5	10
8	7	4	10
9	5	4	10
10	NA	8	NA

```

cw_MATH_Group_proportions = cw_ratings %>%
  group_by(Component,Week,MATH_Group) %>%
  summarise(
    total_marks_for_Group = sum(Marks)
  ) %>%
  # add in the missing "0%" entries
  complete(nesting(Component,Week),MATH_Group, fill = list(total_marks_for_Group = 0)) %>%
  left_join(cw_item_totals) %>%
  ungroup() %>%
  mutate (
    prop_marks = total_marks_for_Group / total_marks_for_paper,
    Week = as.character(Week)
  ) %>%
  bind_rows(
    exam_props
  )

```

```
## `summarise()` regrouping output by 'Component', 'Week' (override with `.groups` argument)
```

```
## Joining, by = c("Component", "Week")
```

```

ILA_components_data = cw_MATH_Group_proportions %>%
  # reorder the Course factor by mean % group A
  mutate(
    gpAmarks = if_else(MATH_Group=="A",prop_marks,0),
    Component = reorder(Component, gpAmarks, mean)
  )

ILA_components_data %>%
  head() %>%
  kable(booktabs = TRUE)

```

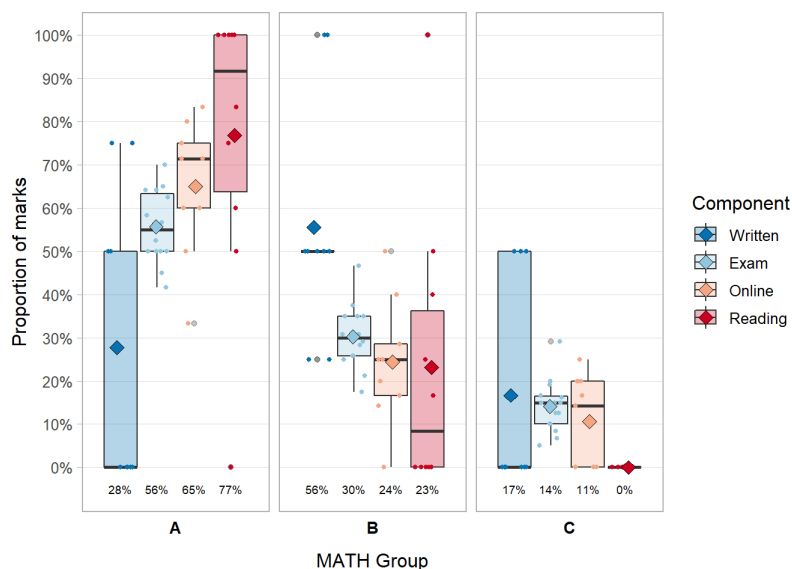
Component	Week	MATH_Group	total_marks_for_Group	total_marks_for_paper	prop_marks	gpAmarks
Online	1	A	3	5	0.6000000	0.6000000
Online	1	B	1	5	0.2000000	0.0000000
Online	1	C	1	5	0.2000000	0.0000000
Online	2	A	5	7	0.7142857	0.7142857
Online	2	B	2	7	0.2857143	0.0000000
Online	2	C	0	7	0.0000000	0.0000000

Plots

```
ILA_components_data %>%
  ggplot(aes(x=1,y=prop_marks, fill=Component)) +
  geom_boxplot(alpha = 0.3) +
  geom_point(aes(colour=Component),
             position = position_jitterdodge(),
             size = 1) +
  stat_summary(fun.y = "mean",
              geom = "point",
              shape = 23,
              size = 3,
              aes(fill = Component),
              position = position_dodge(0.8)) +
  # stat_summary(aes(label=paste0(100*round(..y..,2),"%")),
  #             fun.y=mean,
  #             geom="text",
  #             size=4,
  #             vjust = -0.5,
  #             position = position_dodge(0.75)) +
  geom_text(data = ILA_components_data %>%
            group_by(Component, MATH_Group) %>%
            summarise(prop_marks = mean(prop_marks))),
            aes(label = paste0(100*round(prop_marks,2),"%"),
                y = -0.05),
            position = position_dodge(0.75),
            size = 2.5) +
  scale_y_continuous(labels = scales::percent_format(accuracy = 1),
                    breaks = seq(0,1,0.1)) +
  # scale_fill_brewer(palette = "Set1") +
  # scale_colour_brewer(palette = "Set1") +
  scale_fill_brewer(palette = "RdBu", direction = -1) +
  scale_colour_brewer(palette = "RdBu", direction = -1) +
  theme_light(base_size = 12)+
  theme(
    panel.grid.minor = element_blank(),
    axis.text.x=element_blank(),
    axis.ticks.x=element_blank(),
    panel.grid.major.x = element_blank(),
    strip.background = element_rect(fill=NA,colour = NA),
    strip.text = element_text(size=10, face="bold", color = "black")
  ) +
  labs(
    # title = "Proportion of marks in each MATH Group",
    # subtitle = "Exams are ordered by proportion of Group A marks",
    x = "MATH Group",
    y = "Proportion of marks"
  ) +
  facet_grid( ~ MATH_Group, switch = "x") +
  ggsave("figs/ILA_components_boxplots.pdf",width=15,height=10,units="cm",dpi=300)
```

```
## Warning: `fun.y` is deprecated. Use `fun` instead.
```

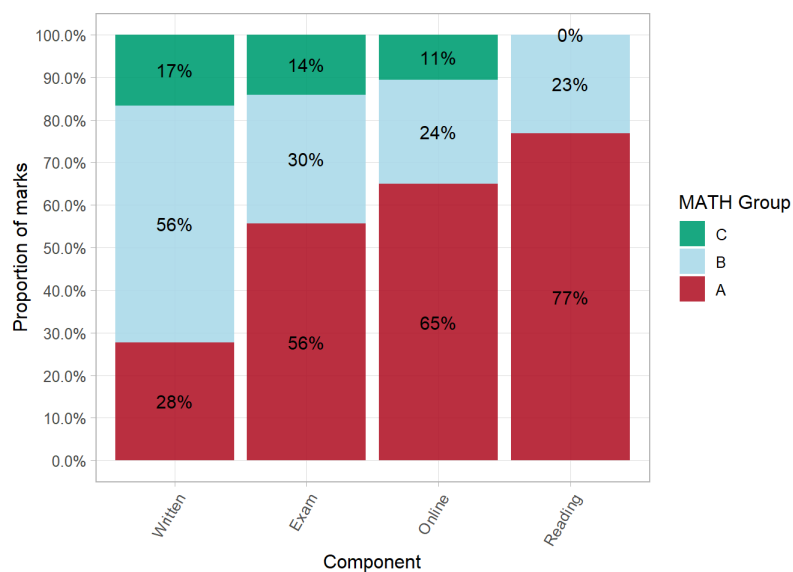
```
## `summarise()` regrouping output by 'Component' (override with `.groups` argument)
```



Using stacked bars:

```
ILA_components_data %>%
  group_by(Component, MATH_Group) %>%
  summarise(prop_marks = mean(prop_marks)) %>%
  # compute the positions of the text labels
  # see http://t-redactyl.io/blog/2016/01/creating-plots-in-r-using-ggplot2-part-4-stacked-bar-plots.html
  ddply(.(Component), transform, pos = cumsum(prop_marks) - (0.5 * prop_marks)) %>%
  ggplot(aes(x = Component,
             y = prop_marks,
             fill = fct_rev(MATH_Group))) +
  geom_bar(stat="identity", alpha = 0.9) +
  geom_text(aes(x = Component, y = pos, label = paste0(sprintf("%.0f", 100*prop_marks), "%")),
            size=4) +
  scale_y_continuous(labels = scales::percent,
                     breaks = seq(0,1,0.1)) +
  scale_fill_manual(values = palMATH) +
  theme_light(base_size = 12)+
  theme(
    panel.grid.minor = element_blank(),
    axis.text.x = element_text(angle = 60, hjust = 1)
  ) +
  labs(
    fill = "MATH Group",
    x = "Component",
    y = "Proportion of marks"
  ) +
  ggsave("figs/ILA_components_stacked.pdf", width=15, height=10, units="cm", dpi=300)
```

```
## `summarise()` regrouping output by 'Component' (override with `.groups` argument)
```



Tables

```
cw_MATH_Cat_proportions = cw_ratings %>%
  group_by(Component, Week, MATH) %>%
  summarise(
    total_marks_for_Cat = sum(Marks)
  ) %>%
  # add in the missing "0%" entries
  complete(nesting(Component, Week), MATH, fill = list(total_marks_for_Cat = 0)) %>%
  ungroup() %>%
  mutate(
    # restore the MATH_Group for the "0%" entries
    Week = as.character(Week),
    MATH_Group = str_sub(MATH, 1, 1)
  ) %>%
  left_join(cw_MATH_Group_proportions) %>%
  mutate (
    prop_marks = total_marks_for_Cat / total_marks_for_Group
  )
```

```
## `summarise()` regrouping output by 'Component', 'Week' (override with `.groups` argument)
```

```
## Joining, by = c("Component", "Week", "MATH_Group")
```

```
ILA_MATH_Cat_props = cw_MATH_Cat_proportions %>%
  group_by(Component, MATH) %>%
  summarise(
    total_Cat = sum(total_marks_for_Cat),
    total_Group = sum(total_marks_for_Group),
    prop_of_Group = total_Cat / total_Group * 100
  )
```

```
## `summarise()` regrouping output by 'Component' (override with `.groups` argument)
```

```
ILA_MATH_Cat_props %>% kable(booktabs = TRUE)
```

Component	MATH	total_Cat	total_Group	prop_of_Group
Online	A1	0.0	32	0.00000
Online	A2	6.0	32	18.75000
Online	A3	26.0	32	81.25000
Online	B1	10.0	12	83.33333
Online	B2	2.0	12	16.66667
Online	C1	1.0	5	20.00000
Online	C2	4.0	5	80.00000
Online	C3	0.0	5	0.00000
Reading	A1	0.0	37	0.00000
Reading	A2	13.0	37	35.13514
Reading	A3	24.0	37	64.86486
Reading	B1	11.0	11	100.00000
Reading	B2	0.0	11	0.00000
Reading	C1	0.0	0	NaN
Reading	C2	0.0	0	NaN
Reading	C3	0.0	0	NaN
Written	A1	0.0	25	0.00000
Written	A2	5.0	25	20.00000
Written	A3	20.0	25	80.00000
Written	B1	7.5	50	15.00000
Written	B2	42.5	50	85.00000
Written	C1	7.5	15	50.00000
Written	C2	7.5	15	50.00000
Written	C3	0.0	15	0.00000

```
ILA_MATH_Cat_props %>%
  select(-total_Cat, -total_Group) %>%
  spread(MATH, prop_of_Group) %>%
  knitr::kable(booktabs = T, digits = 0)
```

Component	A1	A2	A3	B1	B2	C1	C2	C3
Online	0	19	81	83	17	20	80	0
Reading	0	35	65	100	0	NaN	NaN	NaN
Written	0	20	80	15	85	50	50	0

```
ILA_components_data %>%
  mutate(
    Coursework = if_else(Component == "Exam", "Exam", "Coursework")
  ) %>%
  group_by(Coursework, MATH_Group) %>%
  summarise(prop_marks = mean(prop_marks)) %>%
  pivot_wider(
    names_from = Coursework,
    values_from = prop_marks
  ) %>%
  kable(booktabs = TRUE, digits = 2, caption = "Proportion of marks in each MATH Group, for coursework and exams separately.")
```

```
## `summarise()` regrouping output by 'Coursework' (override with `.groups` argument)
```

Proportion of marks in each MATH Group, for coursework and exams separately.

MATH_Group	Coursework	Exam
A	0.57	0.56
B	0.34	0.30
C	0.09	0.14