In-dept Analisys on best Masters of Science in Managemnt worldwide

Francesco Grandi, Tommaso Maccaferri, Cosimo Zatti

RESEARCH QUESTION

This research explores the decision-making process behind choosing the best master's program by integrating economic and career-related factors. The primary question addressed is whether one can assess the quality and value of a master's program by balancing career outcomes with associated costs.

To achieve this, we analyzed data from the Financial Times' *Masters in Management* rankings and combined it with cost-of-living data to contextualize the trade-offs between program prestige, affordability, and long-term career benefits.

Developing our analysis we aimed at answering secondary questions as well, such as what factors most influence the career outcomes of graduates, whether international students prioritize program affordability or prestige when selecting a university, and whether demographic factors like gender and international diversity correlate with post-graduation employment rates.

Web Scraping

We first extract cost-of-living data to combine with information about master's programs.

```r
library(rvest)
library(dplyr)
```

```
Attaching package: 'dplyr'


The following objects are masked from 'package:stats':

    filter, lag
```

```
The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
library(tidyr)
library(tidyverse)
```

```
Warning: package 'ggplot2' was built under R version 4.5.2

-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v forcats   1.0.0     v readr     2.1.5
v ggplot2   4.0.2     v stringr   1.5.1
v lubridate 1.9.4     v tibble    3.3.0
v purrr     1.1.0

-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter()         masks stats::filter()
x readr::guess_encoding() masks rvest::guess_encoding()
x dplyr::lag()            masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becon
```

```
url <- "https://www.numbeo.com/cost-of-living/rankings_by_country.jsp"
page <- read_html(url)

#Extracting the table from the webpage
cost_of_living_table <- page %>%
  html_elements("table") %>%
  html_table(fill = TRUE)
```

Data Cleaning and Preparation

We focus on relevant columns and prepare the cost-of-living data-set for merging with master's program rankings.

```
# Selecting the columns that we need

cost_index_dataframe <- cost_of_living_table[[2]] %>%
  select(`Country` , `Cost of Living Index`, `Rent Index`, `Groceries Index`, `Local Purchas

# Displaying the cleaned dataset
```

```r
print(cost_index_dataframe)
```

```
# A tibble: 155 x 5
   Country           `Cost of Living Index` `Rent Index` `Groceries Index`
   <chr>                              <dbl>        <dbl>             <dbl>
 1 Bermuda                            136.         108.              143.
 2 Cayman Islands                     116.          76.1             124
 3 Us Virgin Islands                  111.          46.8             127.
 4 Switzerland                        111.          51.5             110.
 5 Solomon Islands                    102.          19.5              64.8
 6 Bahamas                             98.8         50.2             101.
 7 Iceland                             97.2         49.5             104.
 8 Jersey                              88.7         52.4              79
 9 Singapore                           87.7         73.1              77.3
10 Norway                              83.7         29.2              85.4
# i 145 more rows
# i 1 more variable: `Local Purchasing Power Index` <dbl>
```

Now we load the Financial Times data-set of master's programs and align it with the cost-of-living data.

```r
# Data-set dowloaded from Financial Times website:
# https://rankings.ft.com/business-education/masters-in-management

library(readxl)

FT_dataframe <- read_excel("/Users/macca/Downloads/export-ranking-masters-in-management-2024-

# Renaming columns for consistency

FT_dataframe_m <- FT_dataframe %>%
  rename(Country = `Location by primary campus`)

# We proceed to eft-join the datasets by "Country", and normalizing country names for compati

FT_dataframe_m$Country <- gsub("US", "USA", FT_dataframe_m$Country)
cost_index_dataframe$Country <- gsub("United States", "USA", cost_index_dataframe$Country)
cost_index_dataframe$Country <- gsub("United Kingdom", "UK", cost_index_dataframe$Country)

# Merging datasets by "Country"
```

```r
merged_data_df<- FT_dataframe_m %>%
  left_join(cost_index_dataframe, by = "Country")

# Displaying the merged dataset

print(merged_data_df)
```

```
# A tibble: 100 x 33
   `#`   `School Name`      Country Weighted salary (US$~1 `Career progress rank`
   <chr> <chr>              <chr>   <chr>                  <chr>
 1 1     University of St~ Switze~ 140,020                60
 2 2     HEC Paris          France  127,375                61
 3 3     Insead             France  118,984                29
 4 4     Edhec Business S~ France  108,239                55
 5 5     Shanghai Jiao To~ China   116,898                6
 6 6     ESCP Business Sc~ France  104,197                58
 7 6     London Business ~ UK      119,823                65
 8 8     EMLyon Business ~ France  102,970                30
 9 8     Nova School of B~ Portug~ 109,874                20
10 10    Essec Business S~ France  111,185                77
# i 90 more rows
# i abbreviated name: 1: `Weighted salary (US$)`
# i 28 more variables: `International work mobility rank` <chr>,
#   `International course experience rank` <chr>, `Rank in 2023` <chr>,
#   `Rank in 2022` <chr>, `Three-year average rank` <chr>,
#   `Programme name` <chr>, `Alumni network rank` <chr>,
#   `ESG and net zero teaching rank` <chr>, `Carbon footprint rank` <chr>, ...
```

Data-set Analysis

Description of the variables of the merged data-set to better orient ourselves in the analysis:

- **Weighted Salary (US$):** Average salary of graduates three years post-completion (PPP-adjusted).
- **Career Progress Rank:** Tracks changes in alumni seniority and organizational size.
- **Value for Money Rank:** Evaluates alumni salary against program costs.
- **Employed at Three Months (%):** Proportion of graduates employed within three months.
- **Cost of Living Index:** Excludes rent; relative to NYC (baseline: 100).
- **Rent Index:** Rent prices relative to NYC.
- **Groceries Index:** Food costs relative to NYC.

- **Local Purchasing Power Index:** Average net salary-based affordability.

To perform calculations, we need to clean and convert string variables to numeric.

```
# Cleaning numeric columns

library(tidyverse)
merged_data_df <- merged_data_df %>%
  mutate(
  "Weighted salary (US$)" = str_remove_all(`Weighted salary (US$)`, ","))

merged_data_df <- merged_data_df %>%
  mutate(
    "Employed at three months (%)" = str_extract(`Employed at three months (%)`, "^[0-9]+")

merged_data_df <- merged_data_df %>%
  mutate(across(
    -c(`School Name`, Country), as.numeric))
```

```
Warning: There were 6 warnings in `mutate()`.
The first warning was:
i In argument: `across(-c(`School Name`, Country), as.numeric)`.
Caused by warning:
! NAs introduced by coercion
i Run `dplyr::last_dplyr_warnings()` to see the 5 remaining warnings.
```

```
# Viewing cleaned data

print(merged_data_df)
```

```
# A tibble: 100 x 33
     `#` `School Name`     Country Weighted salary (US$~1 `Career progress rank`
   <dbl> <chr>             <chr>                    <dbl>                  <dbl>
 1     1 University of St~ Switze~                 140020                     60
 2     2 HEC Paris         France                  127375                     61
 3     3 Insead            France                  118984                     29
 4     4 Edhec Business S~ France                  108239                     55
 5     5 Shanghai Jiao To~ China                   116898                      6
 6     6 ESCP Business Sc~ France                  104197                     58
 7     6 London Business ~ UK                      119823                     65
 8     8 EMLyon Business ~ France                  102970                     30
 9     8 Nova School of B~ Portug~                 109874                     20
```

```
10    10 Essec Business S~ France                   111185                 77
# i 90 more rows
# i abbreviated name: 1: `Weighted salary (US$)`
# i 28 more variables: `International work mobility rank` <dbl>,
#   `International course experience rank` <dbl>, `Rank in 2023` <dbl>,
#   `Rank in 2022` <dbl>, `Three-year average rank` <dbl>,
#   `Programme name` <dbl>, `Alumni network rank` <dbl>,
#   `ESG and net zero teaching rank` <dbl>, `Carbon footprint rank` <dbl>, ...
```

Correlation Analysis: Master's Quality

We analyze how key metrics correlate to assess the quality of master's programs.

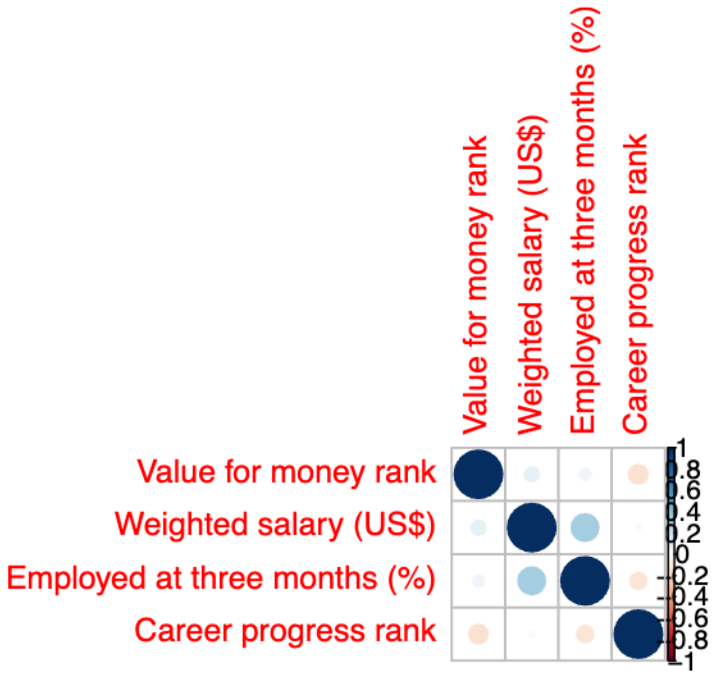
```r
# Correlation matrix for program quality metrics

cor_matrix_quality <- cor(merged_data_df %>%
                            select(`Value for money rank`,
                                   `Weighted salary (US$)`,
                                   `Employed at three months (%)`,
                                   `Career progress rank`))

# Displaying correlation results

print(cor_matrix_quality)
```

```
                             Value for money rank Weighted salary (US$)
Value for money rank                   1.00000000            0.09071148
Weighted salary (US$)                  0.09071148            1.00000000
Employed at three months (%)           0.05475558            0.32451116
Career progress rank                  -0.15428743           -0.01585734
                             Employed at three months (%) Career progress rank
Value for money rank                           0.05475558          -0.15428743
Weighted salary (US$)                          0.32451116          -0.01585734
Employed at three months (%)                   1.00000000          -0.12008937
Career progress rank                          -0.12008937           1.00000000
```

```r
library(corrplot)
```

```
corrplot 0.95 loaded
```

```r
corrplot(cor_matrix_quality)
```



We witness a positive correlation between "Employment at three months" and "Weighted salary". This finding highlights a significant connection between the proportion of graduates employed within three months and their weighted salaries. Programs that effectively place graduates in jobs shortly after completion tend to lead to higher earnings. This could reflect the program's effectiveness in preparing students for lucrative job opportunities or strong industry partnerships.

Moreover, "Value for money" and "Career progress" are negatively correlated. This inverse relationship suggests that programs offering higher value for money (better outcomes relative to cost) may not align with rapid career progress rankings. In contrast, prestigious or expensive institutions with lower perceived value may facilitate faster career advancement, possibly due to their reputation, strong alumni networks, or connections to high-profile industries.

Behavior of International Students

Now we move to investigate factors influencing international student enrollment, including cost indices and program metrics.

```r
# Correlation matrix for international student behavior

cor_matrix_int <- cor(merged_data_df %>%
                        select(`International students (%)`,
                               `Cost of Living Index`,
                               `Rent Index`,
                               `Groceries Index`))

# Correlation with program metrics

cor_matrix_choice <- cor(merged_data_df %>%
                           select(`International students (%)`,
                                  `Value for money rank`,
                                  `Career progress rank`))

print(cor_matrix_int)
```

```
                          International students (%) Cost of Living Index
International students (%)                  1.0000000            0.5352139
Cost of Living Index                        0.5352139            1.0000000
Rent Index                                  0.6325694            0.8559413
Groceries Index                             0.4868324            0.9771450
                          Rent Index Groceries Index
International students (%)  0.6325694       0.4868324
Cost of Living Index        0.8559413       0.9771450
Rent Index                  1.0000000       0.7817197
Groceries Index             0.7817197       1.0000000
```

```r
print(cor_matrix_choice)
```

```
                          International students (%) Value for money rank
International students (%)                 1.00000000          -0.02106979
Value for money rank                      -0.02106979           1.00000000
Career progress rank                      -0.02346692          -0.15428743
                          Career progress rank
International students (%)          -0.02346692
Value for money rank                -0.15428743
Career progress rank                 1.00000000
```

The positive correlation between “Rent costs” and the “Percentage of international students” implies that higher living expenses, often associated with cosmopolitan and attractive urban centers, draw more international applicants. This trend could reflect the preference for studying in prestigious or globally recognized cities, despite their higher costs of living.

On the other hand, the weak correlations between "Value for money rank" and "Career progress rank" with the proportion of international students suggest that these program-specific attributes do not significantly influence international student choices. Instead, factors like the reputation of the location or cultural appeal may play a larger role in their decision-making process.

Employment and Demographic Trends

We proceed to analyze employment rates relative to gender and international student percentages.

Considering this data as been roughly stable through the years, do Universities with more international or female student have higher Employment rates?

```
# Correlation matrix for demographic trends

cor_matrix_emp <- cor(merged_data_df %>%
                        select(`Female students (%)`,
                               `International students (%)`,
                               `Employed at three months (%)`,
                               ))
print(cor_matrix_emp)
```

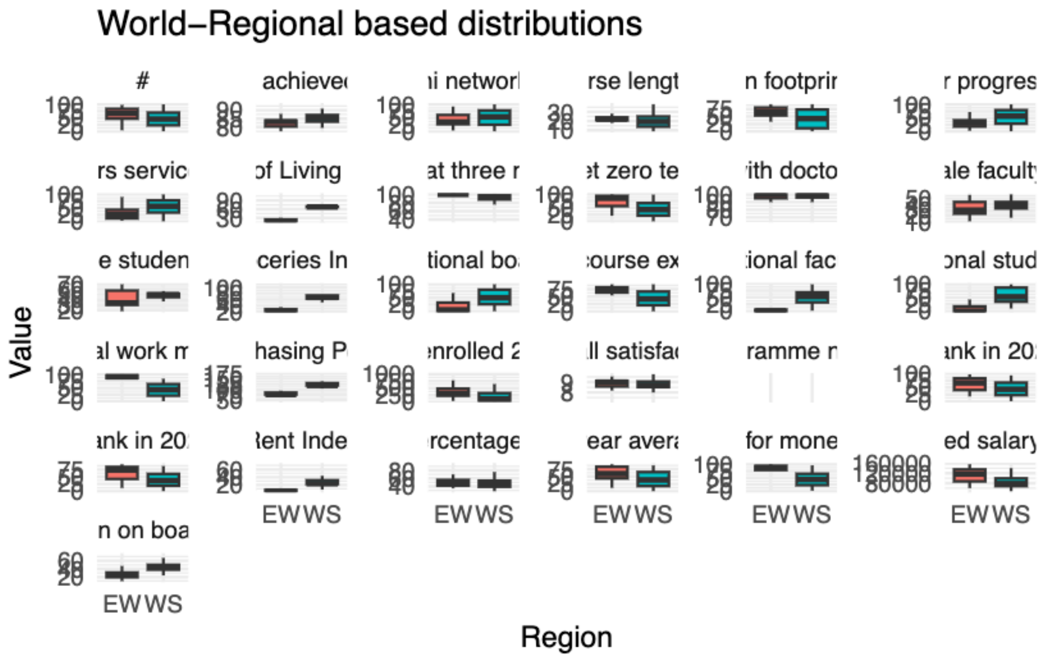
```
                             Female students (%) International students (%)
Female students (%)                    1.0000000                  0.3879720
International students (%)             0.3879720                  1.0000000
Employed at three months (%)          -0.2656603                 -0.1813447
                             Employed at three months (%)
Female students (%)                            -0.2656603
International students (%)                     -0.1813447
Employed at three months (%)                    1.0000000
```

The results show a strong negative correlation between both variables and the percentage of employment. This can suggest two trends: either international students and female are more likely to continue their studying career, justifying why they do not find a job once graduated, or, on the other hand, it can show a tendency for employers to prefer male and national students over the other groups.

Geographical Analysis:

```r
merged_data_df <- merged_data_df %>%
  mutate(Region = case_when(
    Country %in% c("China","Singapore","Australia","India","Taiwan") ~ "EW",
    TRUE ~ "WS"))

merged_data_df <- merged_data_df %>%
  select(`#`,`School Name`, Country, Region, everything())

merged_data_df_long <- merged_data_df %>%
  pivot_longer(cols = -c(Country, Region, `School Name`),
               names_to = "Variable", values_to = "Value")

library(GGally)
merged_data_df_long %>%
  ggplot(aes(x=Region, y=Value, fill=Region))+
  geom_boxplot(outlier.shape = NA) +
  facet_wrap(~ Variable, scales = "free_y") +
  labs(
    title = "World-Regional based distributions",
    x="Region",
    y="Value"
  ) +
  theme_minimal() +
    theme(
      legend.position = "none"
    )
```

```
Warning: Removed 175 rows containing non-finite outside the scale range
(`stat_boxplot()`).
```

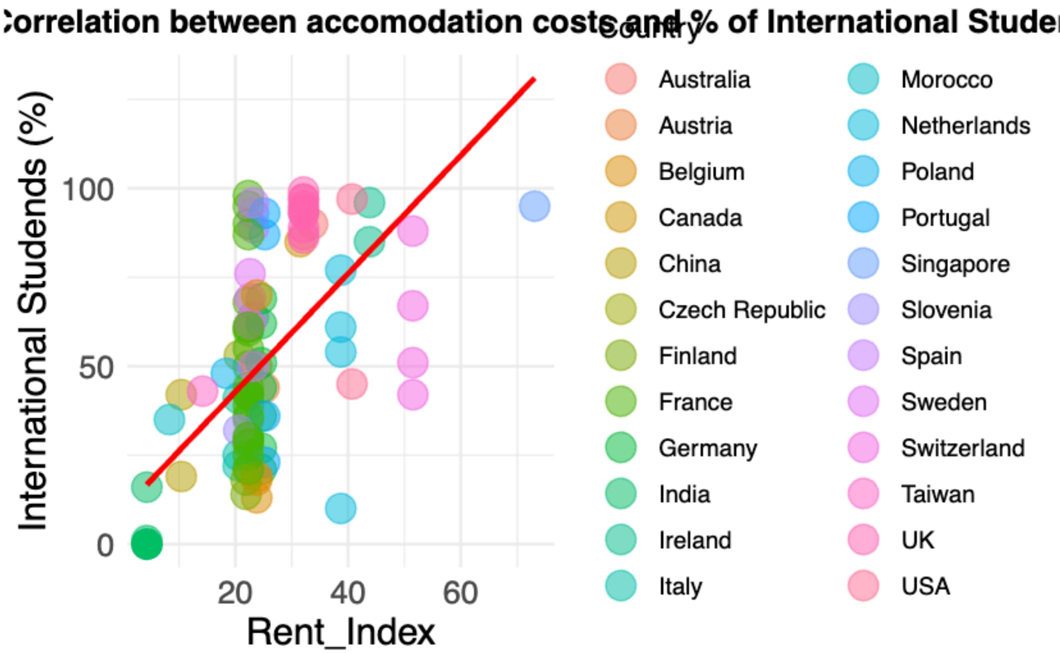


Using graphs these relationships can be better visualized.

Rent Index vs. International Students:

```
library(ggplot2)
cost_int <- ggplot(data=merged_data_df, aes(x=`Rent Index`, y=`International students (%)`, c
           geom_point(size=5, alpha=0.5)+
          geom_smooth(method = "lm", color = "red", se = FALSE) +
           labs(
             x= "Rent_Index",
             y= "International Studends (%)",
             title= "Correlation between accomodation costs and % of International Students"
           ) +
           theme_minimal()+
          theme(
            plot.title = element_text(hjust = 0.2, face = "bold", size = 12),
            axis.text = element_text(size = 12),
            axis.title = element_text(size = 14)
          )

cost_int
```

```
`geom_smooth()` using formula = 'y ~ x'
```



The graph shows a positive correlation between the "Rent Index" and the "Percentage of international students". As the Rent Index increases, the proportion of international students tends to rise.

Observations aligned vertically around a rent index equal to 20 indicate that institutions in countries with similar accommodation costs (similar Rent Index values) have varied proportions of international students. This suggests other factors beyond rent, such as institutional reputation, program offerings, and visa policies, are influencing international student distribution

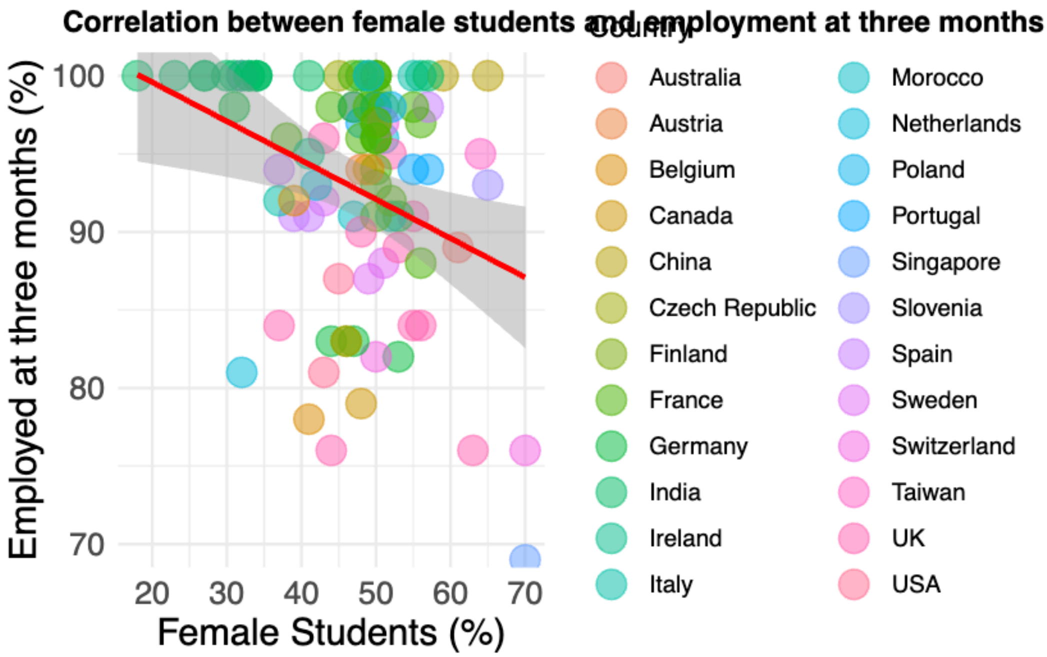
Moreover it is possible to notice that some countries (or institutions) attract high proportions of international students despite moderate or high rent costs. For instance: Institutions with Rent Indexes between 25-30 vary significantly, hosting international student proportions ranging from around 20% to nearly 100%.

Female Students vs. Employment:

```
fem_empl <- ggplot(data=merged_data_df, aes(x=`Female students (%)`, y=`Employed at three mo
  geom_point(size=5, alpha=0.5)+
  geom_smooth(method = "lm", color="red", se=TRUE)+
  labs(
    x="Female Students (%)",
    y="Employed at three months (%)",
    title = "Correlation between female students and employment at three months"
  ) +
  coord_cartesian(ylim = c(70,100)) +
```

```
  theme_minimal()+
theme(
  plot.title = element_text(hjust = 0.1, face = "bold", size = 11),
  axis.text = element_text(size = 12),
  axis.title = element_text(size = 14)
)

fem_empl
```

```
`geom_smooth()` using formula = 'y ~ x'
```



The scatter-plot displays a negative correlation between the "Percentage of female students" and the "Percentage of graduates employed within three months". As the proportion of female students increases, the employment rate at three months tends to decrease.

Since through these plots we are able to highlight only relationships between two variables, the usage of Principal Component Analysis (PCA) is useful for addressing relationships between more than one variable. To this purpose, we proceed to use PCA to visualize the first correlation in the study:

```
# PCA for university quality metrics

library(factoextra)
```

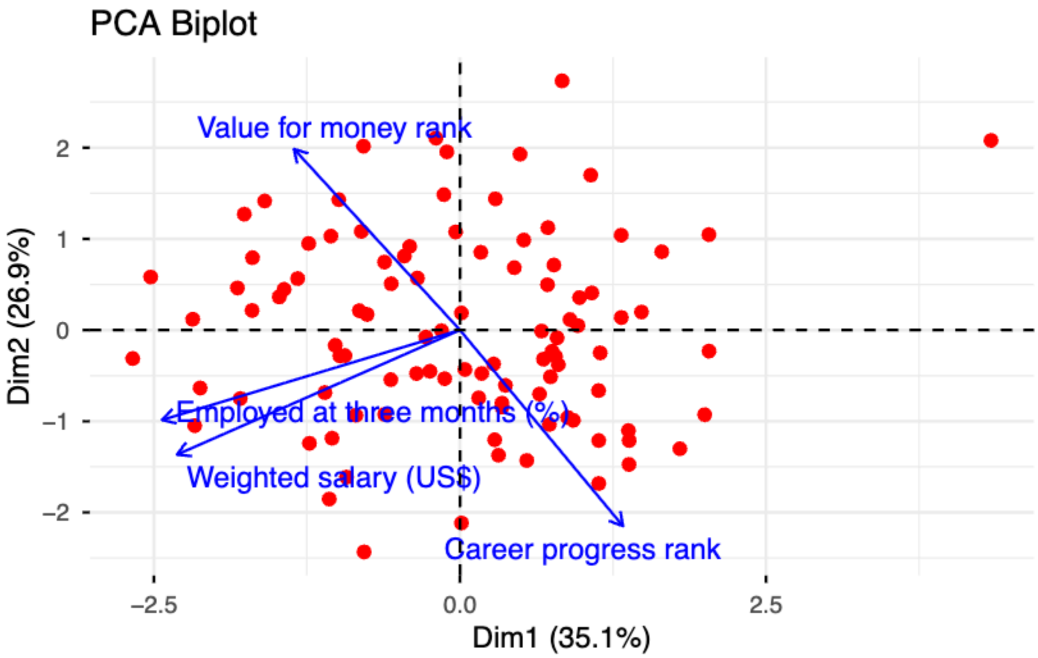
```
Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```r
pca_results <- prcomp(merged_data_df %>%
                        select(`Weighted salary (US$)`, `Career progress rank`, `Employed at
summary(pca_results)
```

```
Importance of components:
                          PC1    PC2    PC3    PC4
Standard deviation     1.1850 1.0380 0.9328 0.8051
Proportion of Variance 0.3511 0.2693 0.2175 0.1620
Cumulative Proportion  0.3511 0.6204 0.8379 1.0000
```

```r
# Visualize the PCA biplot

fviz_pca_biplot(pca_results,
                geom = c("point"),
                pointsize = 2,
                col.var = "blue",
                col.ind = "red",
                title = "PCA Biplot",
                repel = TRUE)
```

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
i The deprecated feature was likely used in the ggpubr package.
  Please report the issue at <https://github.com/kassambara/ggpubr/issues>.
```



This PCA biplot visualizes data with two principal components (Dim1 and Dim2), explaining 62% of the total variance. The red dots represent observations (e.g., universities or programs), while the blue arrows show variables' contributions. Dim1 (35.1%) correlates positively with "Career progress rank," but negatively with "Value for money rank," "Weighted salary (US$)," and "Employed at three months (%)" suggesting a trade-off between career outcomes and affordability. Dim2 (26.9%) distinguishes "Employed at three months (%)" more strongly, highlighting employment as a distinct factor..

Observations near the arrows align with those variables, indicating strong performance in related metrics. For instance, points in the positive Dim1 and Dim2 quadrant likely excel in career progress and employment rates. Patterns suggest that career advancement may come at the cost of value for money.

Conclusion

This research underscores the complexity of selecting a master's program.

We have witnessed how economic considerations and career outcomes can sometimes conflict, as prestigious institutions often accelerate career progress and offer higher post-graduation salaries but are less cost-efficient in terms of "value for money." This trade-off suggests that students seeking rapid professional advancement may need to prioritize reputation and networking opportunities, while those emphasizing financial prudence may opt for less prestigious but more affordable programs.

Moreover, living costs significantly influence decision-making. However, in an apparent contradiction, high living costs in cosmopolitan cities tend to attract more international students, despite reducing overall affordability. This finding implies that non-financial factors, such as lifestyle, global recognition, or the allure of studying in a vibrant urban center, play a critical role in decision-making. For institutions in high-cost areas, this underscores the importance of highlighting their unique cultural, professional, and academic offerings to justify the investment.

Lastly, demographic factors can significantly change outcomes. In fact, the data reveals disparities in employment outcomes based on demographic factors such as gender and international student status. Female graduates and international students appear to face unique challenges, as reflected in lower immediate employment rates. These findings may suggest a need for universities to foster more inclusive environments and strengthen support systems, such as mentorship programs, career services, and employer outreach initiatives, to bridge these gaps and ensure equitable opportunities for all students.

Despite our findings we would like to point out the potential flaws of our analysis. In fact, the study relies heavily on secondary data sources, including rankings and cost-of-living indices, which may not fully capture the nuanced realities of each program or location. For instance, the Financial Times rankings might overemphasize certain metrics like salary while neglecting others like teaching quality or student satisfaction. Moreover, the dataset lacks detailed demographic information on factors such as socioeconomic background, which could influence affordability considerations and career outcomes.

Lastly, it is important to point out that the correlations we identified between variables (e.g., rent index and international students), do not necessarily imply causation. For example, the positive correlation between rent costs and international student enrollment could be influenced by unobserved factors such as visa policies or scholarships.

We believe future research could delve deeper into the qualitative factors influencing program choices, such as cultural fit or alumni networks, to build an even more comprehensive decision-making framework.