Mubanga Nsofu, July 7th 2024

mnsofu@learner.nexford.org

Abstract

A brief description of how to use the R script file to generate payment slip file for Highridge Construction Company

README FILE FOR R PAYSLIP SCRIPT GENERATOR

Contents

[1.0 Employee Payslip Data Generation 2](#_Toc171271388)

[2.0 Requirements 2](#_Toc171271389)

[3.0 Setup 3](#_Toc171271390)

[4.0 Load Libraries Using Package Manager (Pacman): 3](#_Toc171271391)

[5.0 Import Python's faker Python Library into R Using reticulate package: 4](#_Toc171271392)

[6.0 Setup the directories where to store data 4](#_Toc171271393)

[7.0 Generate Employee Data 5](#_Toc171271394)

[8.0 Bugs and Help 7](#_Toc171271395)

[9. Reproducibility 7](#_Toc171271396)

[11. Exception handling 8](#_Toc171271397)

[12. References: 8](#_Toc171271398)

# 1.0 Employee Payslip Data Generation

This R code empowers you to generate synthetic employee payslip data using the Faker library from Python, integrated into R with the help of the reticulate package. The data includes employee ID, name, gender, age, address, salary, overtime hours, leave days, and employee level. The solution is crafted in R studio but can be replicated in other environments such as VS Code, etc.

# 2.0 Requirements

* R version 4.0 or later
* R studio **2024.04.2** Build 764 or later, as shown below



Figure RStudio version

**R packages:**

* pacman
* remotes
* renv
* reticulate
* tidyverse
* Python with faker library

# 3.0 Setup

Install package manager (pacman) and reticulate package, if necessary, as shown below in this piece of code:

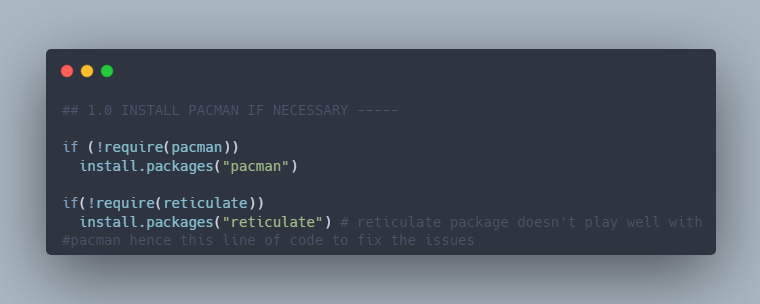


Figure Install pacman and reticulate

Pacman doesn’t play well with the reticulate package, so the line of code above aims to minimise any errors when Pacman uses the p\_load function. If you encounter any issues, rest assured that support is available. Please get in touch with me (details are in the bug section, section 8).

# 4.0 Load Libraries Using Package Manager (Pacman):

**A screenshot of a computer

Description automatically generated**

Figure Load Libraries using pacman.

# 5.0 Import Python's faker Python Library into R Using reticulate package:

**A screen shot of a computer program

Description automatically generated**

Figure Import Python faker package into R

# 6.0 Setup the directories where to store data

The figure below shows how to set up the directory and file and where to store the dataset.

Firstly, do not do this (create a newline) as shown below; otherwise, the code will not work.

****

Figure : Directory and File setup- Don’t do this.

Do this instead

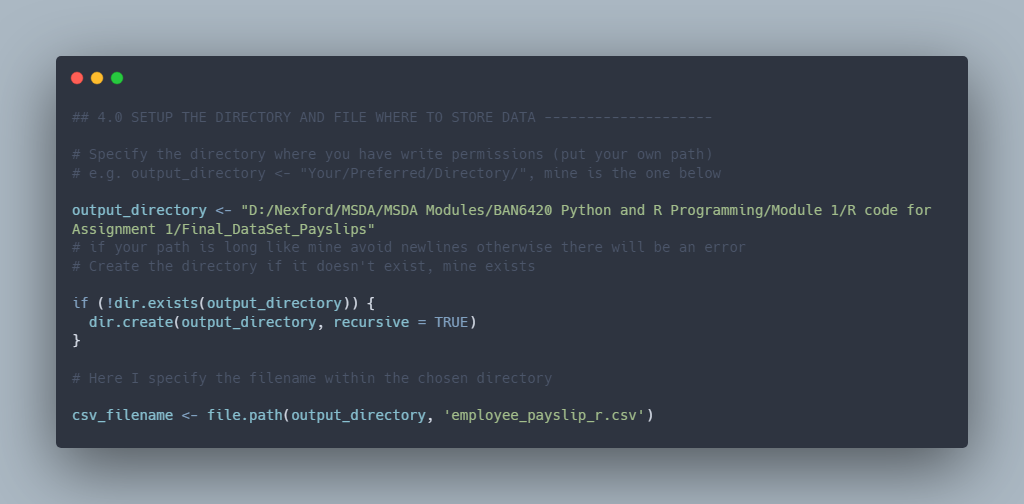
****

Figure : Correct way to specify the paths

# 7.0 Generate Employee Data

Run lines 60 to 138 to generate the employee data. The code will create an output CSV file called ‘employee\_payslip\_r.csv’ in the chosen directory (refer to section 6.0). See the code snippet below.

A screen shot of a computer

Description automatically generated

Figure Code to generate the data

# 8.0 Bugs and Help

In case of any bugs and or help needed, please reach out to me at [mnsofu@learner.nexford.org](mailto:mnsofu@learner.nexford.org)

# 9. Reproducibility

This project uses the renv package to manage dependencies and ensure reproducibility. When you run renv::init() and renv::snapshot(), renv generates a lock file (renv.lock) that records the state of your project's library. This lock file can recreate the same environment later, ensuring that the results are reproducible.

Use the command renv::restore () to restore the environment from the lock file. The text file called README\_lockfile.txt is included for reference. This file provides instructions on how to use the lock file to restore the environment, ensuring reproducibility. The code snippet below illustrates this.

A screen shot of a computer program

Description automatically generated

Figure : Code Reproducibility.

# 11. Exception handling

Exception handling was implemented using tryCatch blocks for each critical section, generating individual employee data and saving the DataFrame to a CSV file. This will ensure that any errors encountered during these processes are captured and reported, allowing the script to continue running without interruption.

# 12. References:

Ushey, K., Allaire, J., & Tang, Y. (2024). reticulate Interface to 'Python' (Version 1.38.0) [R package]. Retrieved from https://github.com/rstudio/reticulate, <https://rstudio.github.io/reticulate/>

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., . . . Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. <https://doi.org/10.21105/joss.01686>

Ushey, K., & Wickham, H. (2024). *renv: Project environments* (Version 1.0.7) [R package]. Retrieved from <https://github.com/rstudio/renv>, <https://rstudio.github.io/renv/>