[](https://in.insidesherpa.com/dashboard)

* Allow ANZ to contact me

[](https://in.insidesherpa.com/virtual-internships/ZLJCsrpkHo9pZBJNY)

[Homepage](https://in.insidesherpa.com/virtual-internships/prototype/ZLJCsrpkHo9pZBJNY/Data%40ANZ%20Program)   >   [Module Overview Page](https://in.insidesherpa.com/virtual-internships/ZLJCsrpkHo9pZBJNY)  >   [Predictive Analytics](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr)

Task 2

**Predictive Analytics**

Explore correlations between customer attributes, build a regression and a decision-tree prediction model based on your findings.

[[](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr#step1)](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step1)

[Watch video instructions](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step1)

[[](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr#step2)](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step2)

[Get background info and context](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step2)

[[](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr#step3)](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step3)

[Get details for your task](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step3)

[[](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr#step4)](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step4)

[View resources to help you learn](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step4)

[[](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr#step5)](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step5)

[Submit your work on this page](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step5)

[[](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr#step6)](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step6)

[Unlock the model work](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BnwCTubtx8NW5W5Kr" \l "step6)

**1**

**Get your instructions from your supervisor**

Watch this video and hear about your task.

Optional: Employment Preferences

**What kind of employment opportunities are you looking for?**

* Pre-internship
* Internships / Vacation schemes
* Full time

🌏  Update working rights  →

📃  Upload your Resume  →

[Download CV/Linkedin Policy](https://insidesherpa-assets.s3-ap-southeast-2.amazonaws.com/files/InsideSherpa+Referencing+Policy.pdf)

This helps companies on InsideSherpa reach out to you for the right roles.

**2**

**Here is the background information on your task**

This task is based on a synthesised transaction dataset containing 3 months’ worth of transactions for 100 hypothetical customers. It contains purchases, recurring transactions, and salary transactions.

The dataset is designed to simulate realistic transaction behaviours that are observed in ANZ’s real transaction data, so many of the insights you can gather from the tasks below will be genuine.

The relevant dataset is linked below for you.

**3**

**Here is your task**

For this task, you’ll likely need to use statistical software such as R, SAS, or Python.

Using the same transaction dataset, identify the annual salary for each customer

Explore correlations between annual salary and various customer attributes (e.g. age). These attributes could be those that are readily available in the data (e.g. age) or those that you construct or derive yourself (e.g. those relating to purchasing behaviour). Visualise any interesting correlations using a scatter plot.

Build a simple regression model to predict the annual salary for each customer using the attributes you identified above

How accurate is your model? Should ANZ use it to segment customers (for whom it does not have this data) into income brackets for reporting purposes?

For a challenge: build a decision-tree based model to predict salary. Does it perform better? How would you accurately test the performance of this model?

**4**

**Resources to help you with the task**

[[](https://insidesherpa.s3.amazonaws.com/vinternships/companyassets/AKkAyEwWc8wjPxx9n/Gz8nAM5drF9tcbBa4/ANZ%20synthesised%20transaction%20dataset.xlsx)](https://insidesherpa.s3.amazonaws.com/vinternships/companyassets/AKkAyEwWc8wjPxx9n/Gz8nAM5drF9tcbBa4/ANZ%20synthesised%20transaction%20dataset.xlsx" \t "_blank)

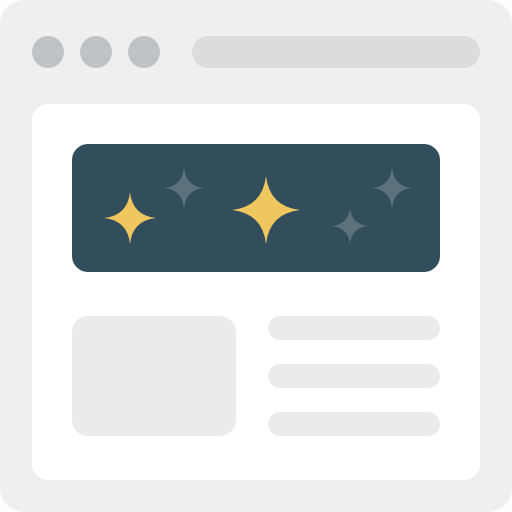
**[ANZ synthesised transaction dataset](https://insidesherpa.s3.amazonaws.com/vinternships/companyassets/AKkAyEwWc8wjPxx9n/Gz8nAM5drF9tcbBa4/ANZ%20synthesised%20transaction%20dataset.xlsx" \t "_blank)**

[Exploratory Data Analysis](https://insidesherpa.s3.amazonaws.com/vinternships/companyassets/AKkAyEwWc8wjPxx9n/Gz8nAM5drF9tcbBa4/ANZ%20synthesised%20transaction%20dataset.xlsx" \t "_blank)

[ANZ synthesised transaction dataset.xlsx](https://insidesherpa.s3.amazonaws.com/vinternships/companyassets/AKkAyEwWc8wjPxx9n/Gz8nAM5drF9tcbBa4/ANZ%20synthesised%20transaction%20dataset.xlsx" \t "_blank)

**[Click to download file →](https://insidesherpa.s3.amazonaws.com/vinternships/companyassets/AKkAyEwWc8wjPxx9n/Gz8nAM5drF9tcbBa4/ANZ%20synthesised%20transaction%20dataset.xlsx" \t "_blank)**

[File copyright ANZ. For use in the Virtual Experience Program only.](https://insidesherpa.s3.amazonaws.com/vinternships/companyassets/AKkAyEwWc8wjPxx9n/Gz8nAM5drF9tcbBa4/ANZ%20synthesised%20transaction%20dataset.xlsx" \t "_blank)

[[](https://www.tableau.com/)](https://www.tableau.com/" \t "_blank)

**[Link to Tableau Website](https://www.tableau.com/" \t "_blank)**

[Exploratory Data Analysis](https://www.tableau.com/" \t "_blank)

[Tableau Website](https://www.tableau.com/" \t "_blank)

**[Click to access website →](https://www.tableau.com/" \t "_blank)**

[File copyright ANZ. For use in the Virtual Experience Program only.](https://www.tableau.com/" \t "_blank)

**For those who would like to try this task using R:**

1. Go to https://notebooks.azure.com/, to create a free R notebook project. After users have uploaded the data, they can create an R project directly in the browser without requiring any installation of software. They can also export or share their work easily to submit it to us.
2. Download and install RStudio by following the instructions at https://www.rstudio.com/

**5**

**Submit your work**

You may only credit this task to your CV if you make a genuine attempt at the work. You may resubmit work, iterate upon and reattempt work as many times as you want.

**Practice or learning attempt:**

 toggle this switch to flag this was a practice attempt to ANZ

Top of Form

Drop files here or click to upload.(You can resubmit your work as many times as you wish)

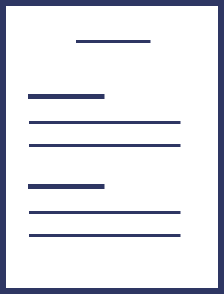
Bottom of Form

**Your completed task attempts:**

**6**

**Unlock the model work**

Download an example of work done at ANZ in this topic. Submit an attempt at work first before this unlocks.



[← Previous Task](https://in.insidesherpa.com/modules/ZLJCsrpkHo9pZBJNY/BiJPfqmGY2QwgN6gA)

[Go to the Module Overview Page](https://in.insidesherpa.com/virtual-internships/ZLJCsrpkHo9pZBJNY)

**What it’s like to work with us…**

Data Science is a contact sport at ANZ.  
  
We look for ways for you to connect with data colleagues internally and externally. We bring together people from across the bank to get value in the hands of customers and colleagues as easily as possible, autonomy, purpose and mastery wrapped together.  
  
Growing our people through development opportunities and work experiences is a critical part of how we evolve and strengthen your capabilities.  
  
Up for the challenge? We’d love to hear from you!

[Explore ANZ Graduate Careers    →](https://www.anz.com.au/careers/programs/graduates/" \t "_blank)

[Explore ANZ Internships    →](https://www.anz.com.au/careers/programs/summer-intern/" \t "_blank)

[](https://in.insidesherpa.com/)

[Privacy Policy](https://in.insidesherpa.com/privacy) [Terms of Use](https://in.insidesherpa.com/terms) © Copyright 2020