



## CSIS 4260 – 002: Assignment 1

### Stock Price Analysis & Prediction

#### Dashboard Screenshots

**Prepared For:**

Nikhil Bhardwaj

**Date:** February 26, 2026

**Github Link:** <https://github.com/desporamon/stock-analysis-a1>

**Live Dashboard:** <https://sp500-desmond-stockprediction.streamlit.app>

**Prepared By:**

Name

Desmond Chua

Student ID

300369803

Email Address

chuad1@student.douglascollege.ca

# Page 1: Overview

The overview page displays key dataset statistics (619,040 rows, 505 companies, date range 2013–2018), a brief project description explaining what each dashboard page covers, and a sample of the data table showing both original columns and the calculated technical indicators.

This screenshot shows the overview page of a dashboard titled "S&P 500 Stock Price Analysis & Prediction". The sidebar on the left contains a "Navigation" section with links to "Overview", "Benchmark Results", "Stock Prediction", and "Technical Indicators". Below the sidebar, it says "CSIS 4260 — Assignment 1" and "Desmond Chua". The main content area has a title "S&P 500 Stock Price Analysis & Prediction" and a subtitle "CSIS 4260 — Assignment 1 | Author: Desmond Chua". It displays summary statistics: Total Records (619,040), Companies (505), Date Range (2013-02-08 to 2018-02-07), and To (2018-02-07). A section titled "About This Dashboard" provides a brief description of the project and lists the four main analysis components. A note at the bottom encourages users to use the sidebar to navigate between pages and select different stock tickers.

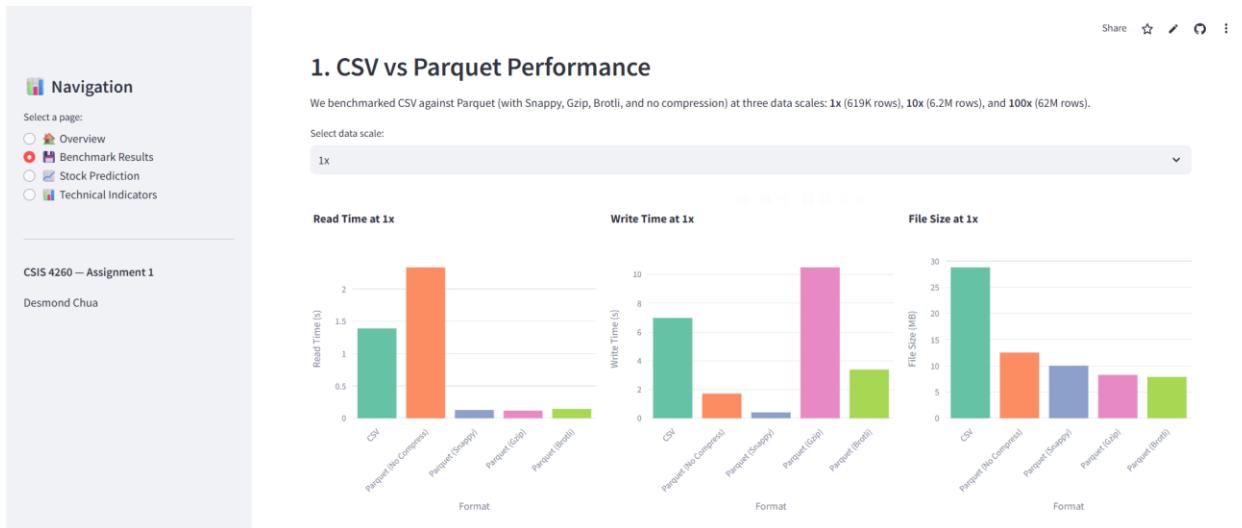
This screenshot shows the "Dataset Sample" page of the dashboard. The sidebar is identical to the overview page. The main content area starts with a note about the sidebar navigation and the availability of "Stock Prediction" and "Technical Indicators" sections. Below this is a table titled "Dataset Sample" showing a subset of the dataset. The table has 16 columns: date, open, high, low, close, volume, name, sma\_20, sma\_50, rsi\_14, volatility, daily\_return, price\_momentum, and next\_close. The data consists of 10 rows, each representing a specific date and time point for a company named 'A'.

	date	open	high	low	close	volume	name	sma_20	sma_50	rsi_14	volatility	daily_return	price_momentum	next_close
0	2013-02-08 00:00:00	45.07	45.35	45	45.08	1,824,755	A	45.08	45.08	50	None	None	0	44.6
1	2013-02-11 00:00:00	45.17	45.18	44.45	44.6	2,915,405	A	44.84	44.84	0	0.3394	-1.0648	-0.24	44.62
2	2013-02-12 00:00:00	44.81	44.95	44.5	44.62	2,373,731	A	44.7667	44.7667	4	0.2715	0.0448	-0.1467	44.75
3	2013-02-13 00:00:00	44.81	45.24	44.68	44.75	2,052,338	A	44.7625	44.7625	23.8095	0.2219	0.2913	-0.0125	44.58
4	2013-02-14 00:00:00	44.72	44.78	44.36	44.58	3,826,245	A	44.726	44.726	18.75	0.2088	-0.3799	-0.146	42.25
5	2013-02-15 00:00:00	43.48	44.24	42.21	42.25	14,657,315	A	44.3133	44.3133	4.7923	1.0279	-5.2266	-2.0633	43.01
6	2013-02-19 00:00:00	42.21	43.12	42.21	43.01	4,116,141	A	44.1271	44.1271	23.3933	1.0598	1.7988	-1.1171	42.24
7	2013-02-20 00:00:00	42.84	42.85	42.225	42.24	3,873,183	A	43.8913	43.8913	19.5279	1.1865	-1.7903	-1.6512	41.63
8	2013-02-21 00:00:00	42.14	42.14	41.47	41.63	3,415,149	A	43.64	43.64	17.2676	1.3417	-1.4441	-2.01	41.8
9	2013-02-22 00:00:00	41.83	42.07	41.58	41.8	3,354,862	A	43.456	43.456	19.8529	1.3923	0.4084	-1.656	41.29

# Page 2: Benchmark Results

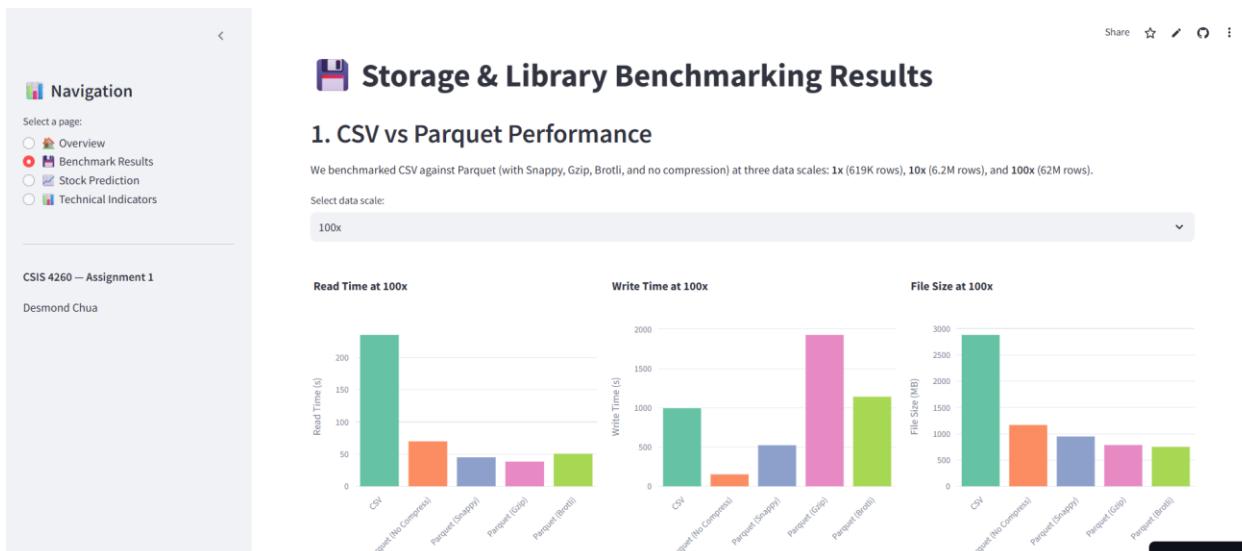
## CSV vs Parquet — 1x Scale

Bar charts comparing read time, write time, and file size for CSV and four Parquet variants (no compression, Snappy, Gzip, Brotli) at the original data scale of 619K rows.



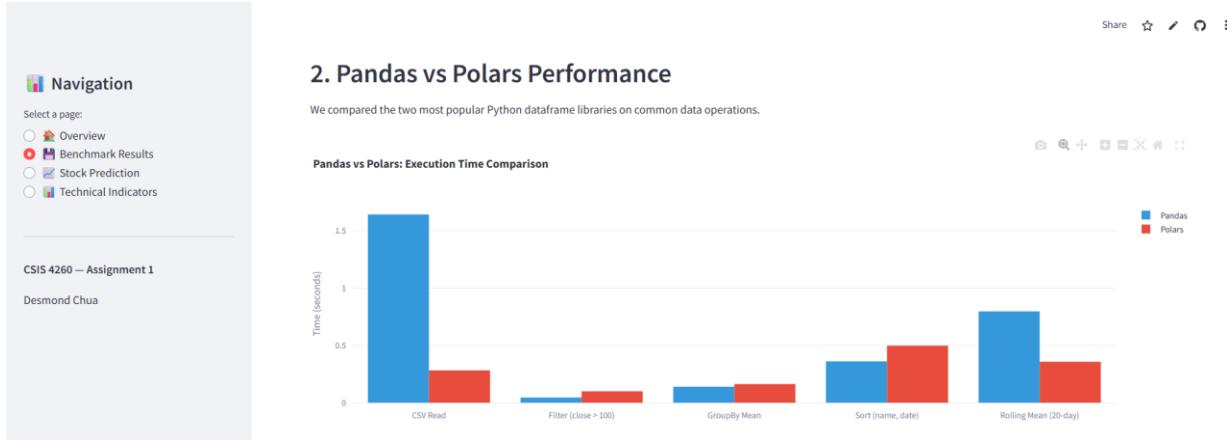
## CSV vs Parquet — 100x Scale

Same charts at 100x scale (62M rows). At this scale the performance gap is dramatic — CSV takes almost 4 minutes to read while Parquet Snappy takes 45 seconds.



## Pandas vs Polars Comparison

Grouped bar chart and results table comparing execution times for Pandas and Polars across five common data operations. Polars is faster for CSV reading (5.8x) and rolling calculations (2.2x), but Pandas was chosen for this project due to better Scikit-learn and Streamlit integration.



# Page 3: Stock Prediction

## Company Ticker Selection

The dashboard provides a searchable dropdown containing all 505 S&P 500 company tickers. Selecting a different company updates all prediction charts and error analysis below in real time. This demonstrates the dashboard meets the requirement of making predictions available for all companies.

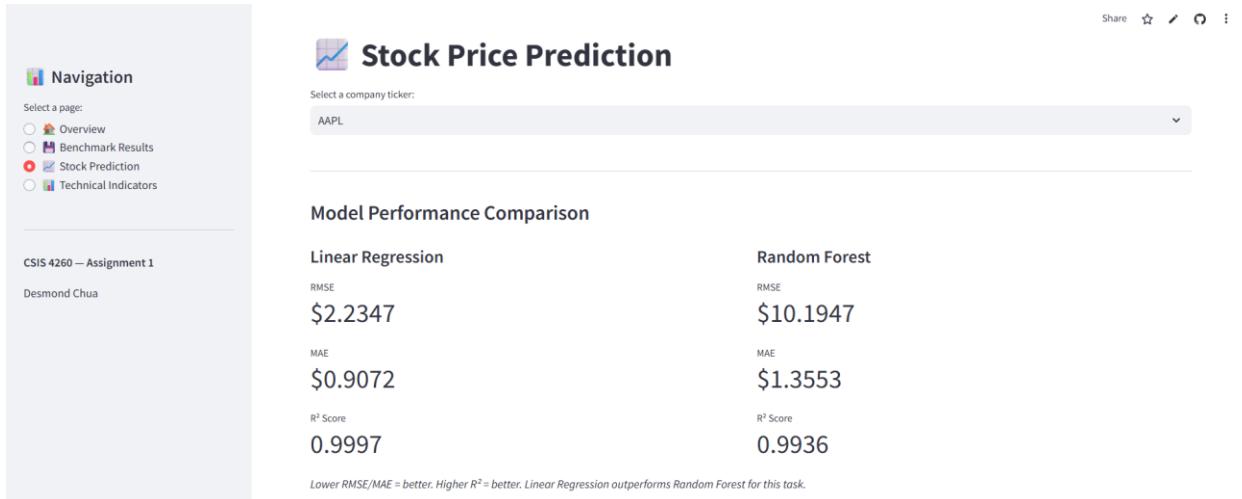
The screenshot shows a dashboard titled "Stock Price Prediction". On the left, there is a "Navigation" sidebar with links to Overview, Benchmark Results, Stock Prediction (which is selected and highlighted in red), and Technical Indicators. Below the sidebar, it says "CSIS 4260 — Assignment 1" and "Desmond Chua". The main content area has a title "Stock Price Prediction" with a chart icon. Below the title is a dropdown menu labeled "Select a company ticker:" containing the text "AJG". A list of company tickers is displayed: A, AAL, AAP, AAPL, ABBV, ABC, ABT, and ACM. To the right of the dropdown, there are two sets of numerical values: \$0.9072 and 0.9997 on the left, and \$1.3553 and 0.9936 on the right. Below these values is a note: "Lower RMSE/MAE = better. Higher R<sup>2</sup> = better. Linear Regression outperforms Random Forest for this task."

	Model 1	Model 2
Value	\$0.9072	\$1.3553
R <sup>2</sup> Score	0.9997	0.9936

Lower RMSE/MAE = better. Higher R<sup>2</sup> = better. Linear Regression outperforms Random Forest for this task.

## Model Performance Metrics

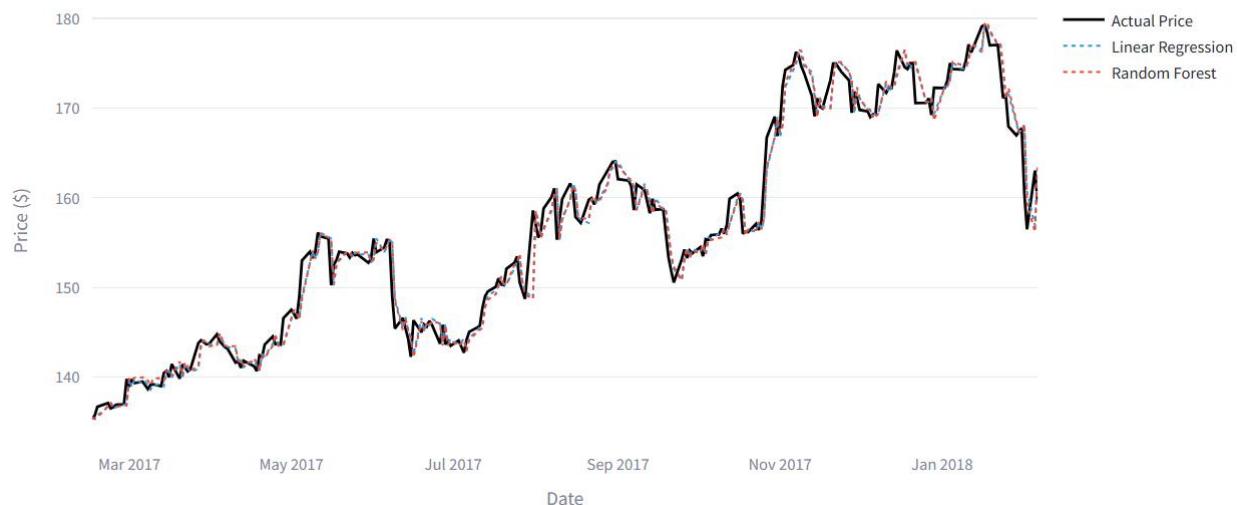
Side-by-side comparison of Linear Regression and Random Forest performance. Linear Regression outperformed across all metrics (RMSE: \$2.23 vs \$10.19, MAE: \$0.91 vs \$1.36, R<sup>2</sup>: 0.9997 vs 0.9936). This is because stock prices follow a near-linear day-to-day pattern.



## Actual vs Predicted — AAPL

Actual closing price (black line) compared with Linear Regression (blue dotted) and Random Forest (red dotted) predictions for Apple on the test set. Both models closely track the actual price.

**AAPL — Actual vs Predicted Closing Price (Test Set)**



## Prediction Error — AAPL

Error chart showing how far each model's prediction deviates from the actual price over time. Linear Regression error stays near \$0, while Random Forest shows larger spikes.

**AAPL — Prediction Error Over Time**



## Actual vs Predicted — GOOG

Same prediction view for Google, demonstrating that the company ticker dropdown works and all charts update dynamically when a different company is selected.

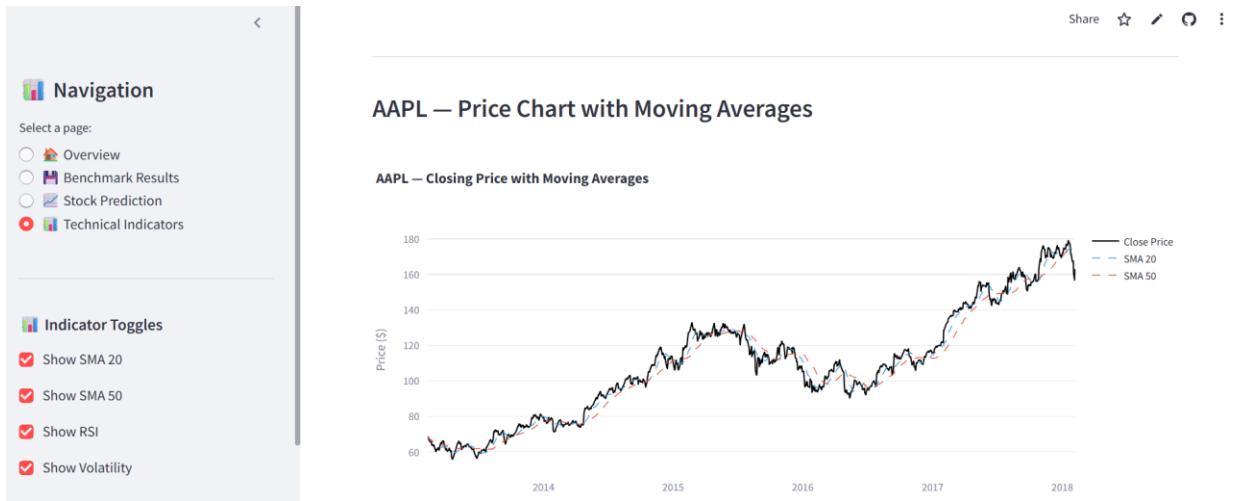
**GOOG — Actual vs Predicted Closing Price (Test Set)**



## Page 4: Technical Indicators

### Price Chart with Moving Averages

Closing price line chart with SMA 20 (short-term) and SMA 50 (medium-term) moving average overlays. The sidebar shows toggle checkboxes that allow turning each indicator on and off interactively.

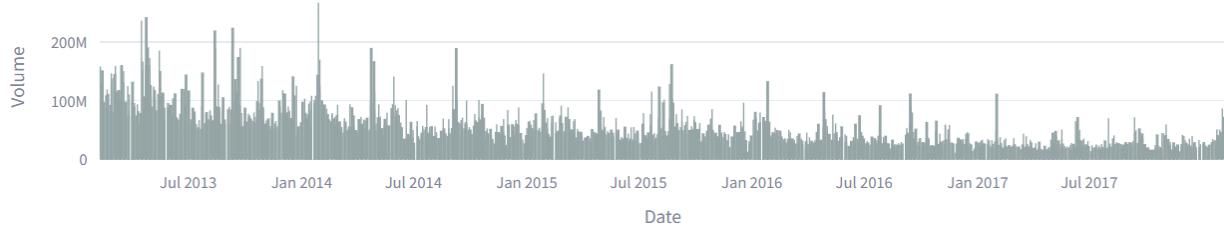


## Trading Volume

Daily trading volume bar chart for the selected company, showing how actively the stock was traded each day.

### AAPL — Trading Volume

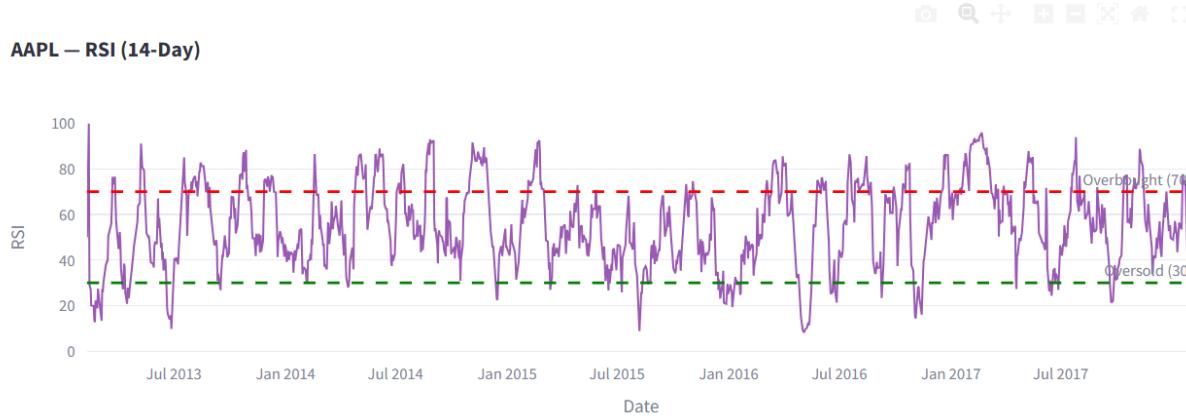
**AAPL — Daily Trading Volume**



## RSI (Relative Strength Index)

RSI chart with overbought (70) and oversold (30) threshold lines marked. RSI measures momentum — values above 70 suggest the stock may be overpriced, below 30 suggests it may be undervalued.

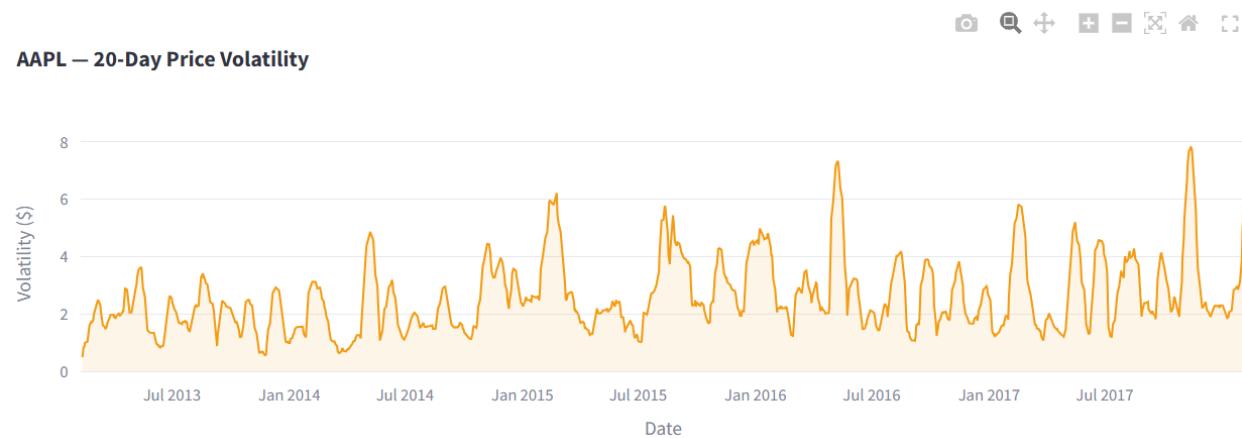
### AAPL — Relative Strength Index (RSI)



## Volatility

20-day rolling standard deviation chart showing price volatility over time. Spikes indicate periods where the stock price was swinging significantly.

### AAPL — 20-Day Volatility



## Summary Statistics

Descriptive statistics table for the selected company, including count, mean, standard deviation, min, max, and quartile values for close price, volume, SMA, RSI, volatility, and daily return.

### AAPL — Summary Statistics

	close	volume	sma_20	rsi_14	volatility	daily_return
count	1,259	1,259	1,259	1,259	1,258	1,258
mean	109.07	54,047,899.74	108.31	55.01	2.61	0.08
std	30.56	33,468,353.34	30.21	18.45	1.27	1.46
min	55.79	11,475,922	59.38	8.21	0.5	-7.99
25%	84.83	29,694,376.5	80.64	41.94	1.69	-0.64
50%	109.01	45,668,931	109.18	54.16	2.29	0.04
75%	127.12	68,708,720	126.48	69.5	3.23	0.86
max	179.26	266,833,581	174.6	100	7.81	8.2