People x Expertise x Data. Boom.

For those driven to accelerate the pace of transformation, RapidMiner is the enterprise-ready data science platform that amplifies the collective impact of your people, expertise and data for breakthrough competitive advantage.
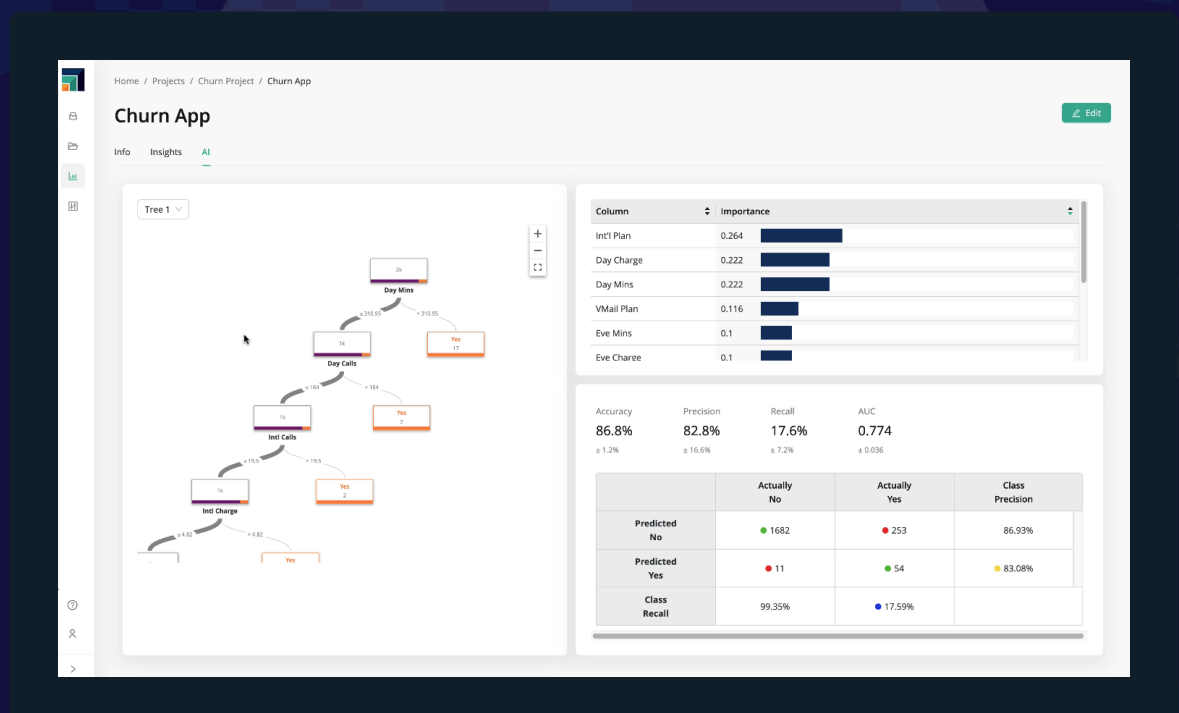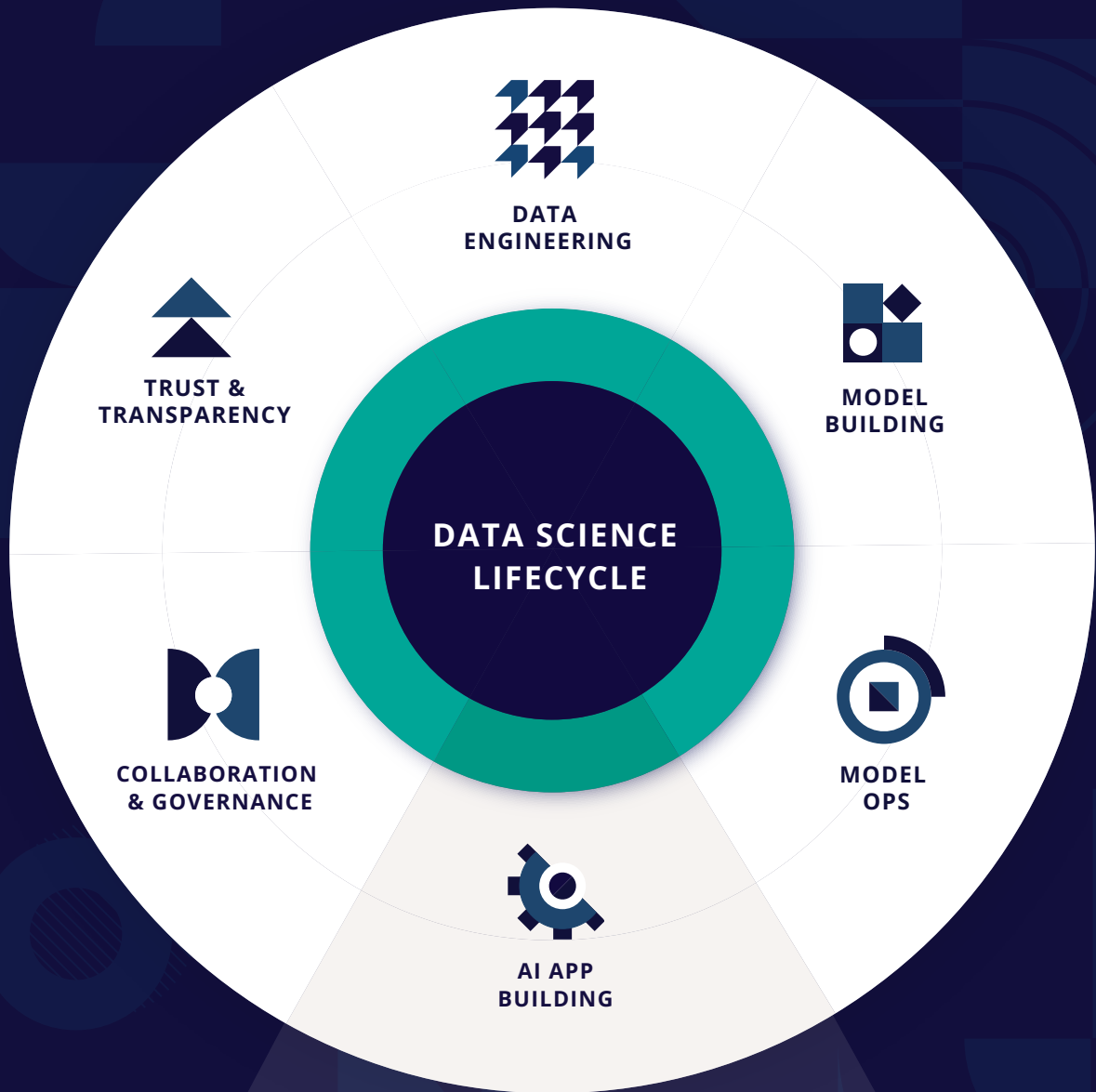
RAPIDMINER

An △ ALTAIR Company

RAPIDMINER
An △ ALTAIR Company

Say Hello to Our New Enterprise AI Platform

Now in Public Beta, RapidMiner's next-gen platform is purpose-built to help you overcome the hurdles that get in the way of Enterprise AI. Beneath a sleek and intuitive exterior, you'll find a powerful analytics engine that enhances the productivity for your entire team. The platform offers one place where anyone can do real data science regardless of their skill level, transparency across the entire AI lifecycle to build trust in solutions, and the flexibility to fit in with your data strategy today, tomorrow, and in the future.

Here's how your enterprise will benefit:

- **One platform for everyone:** Successful data science projects rely on business and data experts to work together. RapidMiner offers full automation for data science novices, an integrated JupyterLab environment for experts, and a visual workflow designer to bridge the gap—all packaged in a single browser-based experience.

- **True team transparency:** Deliver complete transparency across every stage of the AI lifecycle. RapidMiner acts as your single source of truth by centrally storing projects , features, models and data assets, and helps you provide tangible, understandable, and actionable results to decision-makers within your organization.

- **Digital enterprise agility:** Don't let IT challenges get in the way of AI solutions. Our full-SaaS cloud solution let's you add users with ease, scale on demand, and embed models wherever they can deliver the greatest impact for your business—all without compromising the security or integrity of your data.

DATA
ENGINEERING

TRUST &
TRANSPARENCY

MODEL
BUILDING

DATA SCIENCE
LIFECYCLE

COLLABORATION
& GOVERNANCE

MODEL
OPS

AI APP
BUILDING

Churn App

✎ Edit

Info   Insights   AI

Tree 1 ∨

+
−
⛶

1k
Day Mins

≤ 310.95   > 310.95

1k
Day Calls

Yes
17

≤ 164   > 164

1k
Intl Calls

Yes
2

≤ 19.5   > 19.5

1k
Intl Charge

Yes
2

≤ 4.82   > 4.82

Yes

Column ⇅	Importance ⇅
Int'l Plan	0.264
Day Charge	0.222
Day Mins	0.222
VMail Plan	0.116
Eve Mins	0.1
Eve Charge	0.1

Accuracy	Precision	Recall	AUC
86.8%	82.8%	17.6%	0.774
± 1.2%	± 16.6%	± 7.2%	± 0.036

	Actually No	Actually Yes	Class Precision
Predicted No	1682	253	86.93%
Predicted Yes	11	54	83.08%
Class Recall	99.35%	17.59%	

Data Engineering

Spend less time on data "grunt work" and more time building and operationalizing business-changing solutions. RapidMiner gives you everything you need to connect, acquire, explore, prepare, catalog and pipeline your data.

Key features

- Support for NoSQL databases MongoDB & Cassandra

- Access to time series data, audio files, images, and much more

- Feature recommendations, automatic engineering & sharing

- 30+ built-in data visualizations

- Guided data prep

Connect to diverse data sources

Project success relies on being able to connect to relevant data wherever it lives. That's why RapidMiner allows you to connect to all common enterprise applications, read from over 40 file types, and ingest any type of data no matter it's volume, velocity, or variety.

Visualize & explore data

Thoroughly understand your data and identify impactful areas for analysis. RapidMiner offers over 30 interactive visualizations to identify patterns, trends, distributions, and quality issues within your data.

Simplify data prep

Fast, and interactive data prep helps you ensure that your data is always analytics-ready. Merge data from multiple sources, fix quality issues, and save any data prep or ETL pipeline for automation and future use within your organization.

Automate feature selection

RapidMiner's automated feature engineering allows you to generate and select the best attributes for the problem you're trying to solve and ensure that models are their most accurate and understandable based on what's worked for your peers.

Model Building

3 unique, immersive & interchangeable experiences allow you to streamline model creation—whether you're new to data science or a seasoned pro.

Key features

- Automated, visual, and code-based interfaces

- 1500+ built-in algorithms and functions

- Pre-built templates for churn, fraud detection, predictive maintenance and more

- Plain text explanations and guided experience within templates

- Proven validation techniques like cross and split validation

Provide an interface for everyone

Project contributors can choose between RapidMiner's code-based, visual, and automated model creation experiences depending on their skills and preference. Regardless of where you build, all your work will automatically be logged back to a visual workflow for full transparency.

Code-like flexibility, without the complexity

Visual doesn't have to mean limited. RapidMiner has over 1500 algorithms and functions that aren't just designed to replicate the latest machine learning techniques, but also replicate code-like tasks such as looping and branching.

Get to results faster

RapidMiner helps you work smarter, not harder. Pre-built templates for common use cases and proactive recommendations based on what's worked for your peers help to speed up model creation and accelerate time to value.

Check your work

RapidMiner helps you thoroughly validate your models before making the case for deployment. Accurately estimate model performance and control model overfitting by using tried and true techniques like cross and split validation.

Model Ops

Simple, intuitive operation of any model, built by any user. Benefit from code-free deployment, automatic monitoring, bias detection and streamlined insight delivery.

Key features

- Containerized architecture for code and hassle-free deployment
- Value sensitive scoring to quantify business impact
- Automatic monitoring and drift prevention
- Clear & actionable insight delivery

Eliminate handoff friction

RapidMiner helps you create a smooth handoff between model design and production. Use the platform's containerized architecture and code-free model ops to avoid deployment friction and operationalize models so they can make their most valuable predictions.

Demonstrate accuracy AND impact

When you're making the case for deployment, lead with true business impact, not just accuracy. Use value-sensitive scoring to quantify the financial impact of your models—increases in revenue, decreases in costs, and overall impact to the bottom line.

Prevent degradation & drift

Ensure continual value delivery without relying on data scientists to manage your active models. Use RapidMiner to automatically monitor for drift, performance degradation and service health. Identify problematic trends and address them before they start to negatively impact projects.

AI App Building

Put models and insights into the hands of decision-makers and action-takers with a no-code approach.

Key features

- No-code dynamic app creation
- Model simulations and "what if" scenarios
- Fast & easy results sharing
- Inclusive of any project content

Share stories, not just results

RapidMiner simplifies AI app creation so you can show off your work without relying on software engineers. Apps work with any model, whether they're created with autoML, visual workflows, code, or a combination of the three. Tell the full story around your data pipeline and make results easy to understand and act on.

Build trust through engagement

AI Apps enable you to build trust in predictions in a hands-on way. By exposing simulations, evaluating results against expectations, and running what-if scenarios, you make results more approachable for decision-makers who weren't involved in model creation.

Power smarter decision-making

Apps can beyond exposing predictions and analytics results to showcase any and all of your data science project content. Enable model consumers to engage with your work in a variety of ways across any of their devices.

Collaboration & Governance

Enable open communication, easy sharing, and widespread re-use of work without breaking policies and regulations.

Key features

- Project-based framework
- Automatic process explanations
- Real-time collaboration within visual workflows
- SSO, 2FA, TLS support, versioning and rollback capabilities etc.

Work in a transparent environment

RapidMiner's project-based framework allows you to organize all of your work by use case, and easily share it with others in your organization when they're joining an initiative or working on a similar problem.

Create a shared language

Any work that's completed in RapidMiner's code-based or automated experiences is automatically logged back to visual workflows to maintain a universal language for projects. Team members can also co-edit processes within the designer in real-time.

Establish full control

RapidMiner powers data-driven transformation without introducing unnecessary risk. The platform's 4-layer security framework gives admins the ability to authenticate and authorize users, while encrypting enterprise data and auditing it over time.

Trust & Transparency

Instill confidence in your work with explainable AI.

Key features

- Detailed visual process explanations

- Global & local feature weights

- Shareable model simulators

Get the full picture

RapidMiner's visual workflows enable you to see the big picture through explainable analytics. Fully understand your data processing pipeline, review detailed explanations of every modeling step, and see how data is being processed and scored.

Leverage proven explainability tools

Calculate global and local feature weights to understand driving factors for predictions and examine which features are most influential. Explore partial dependencies to see how models interact with a single column in a dataset.

Instill confidence in models

Stakeholders can't always spare the time and effort to learn advanced methods. No-code app creation allows your team to easily visualize what your models are doing and intuitively understand how they work.

Automated Data Science

Full lifecycle support and guidance for novices, productivity enhancements for experts

Fast & intuitive prep

Interactively explore data to evaluate and understand its health, completeness, and quality

Less time to insight

Use code-free modeling and ops to create and deploy impactful models in a matter of clicks

Data scientist in a box

Automate under-considered parts of the DS lifecycle like risk assessment and target recommendations

Visual Workflows

Drag-and-drop model creation experience with the power and flexibility of code

Central source of truth

Get a complete view of your processes down to the step so you can troubleshoot easily and show your work

All the power of code

Use visual building blocks to create models and replicate coding concepts like looping and branching

Ready for customization

Unpack any visual building block within the designer for further tuning, inspection, and customization

Code-Based Data Science

Embedded JupyterLab environment for advanced data science work

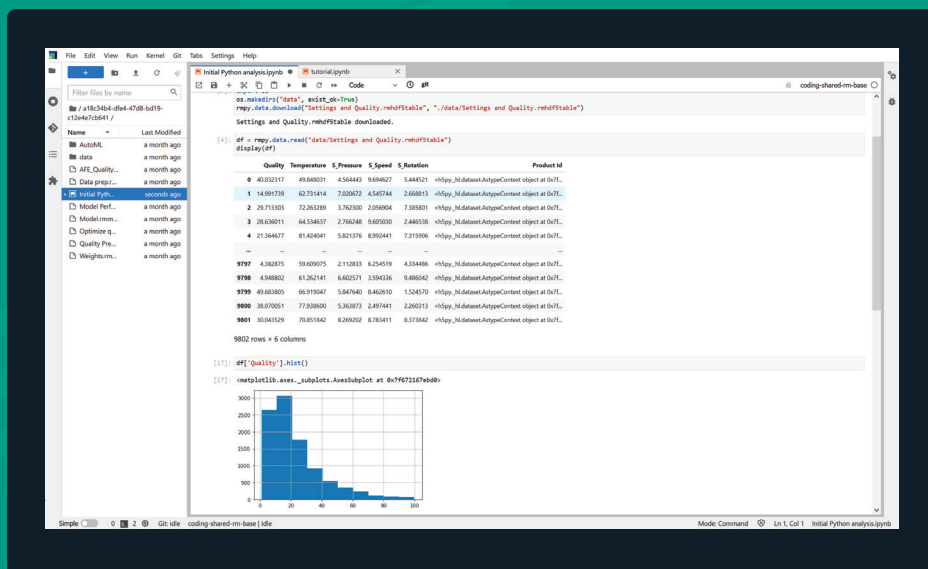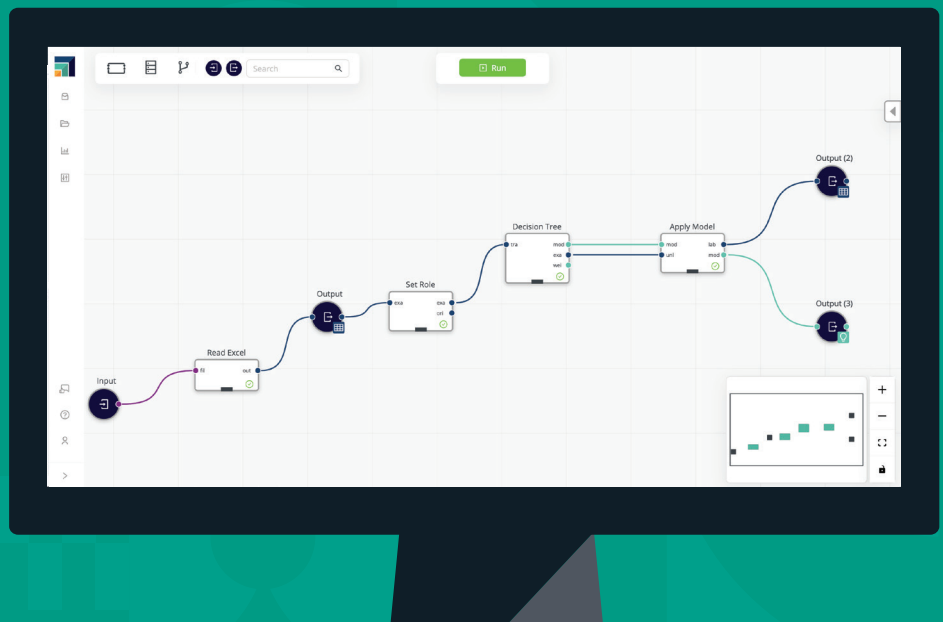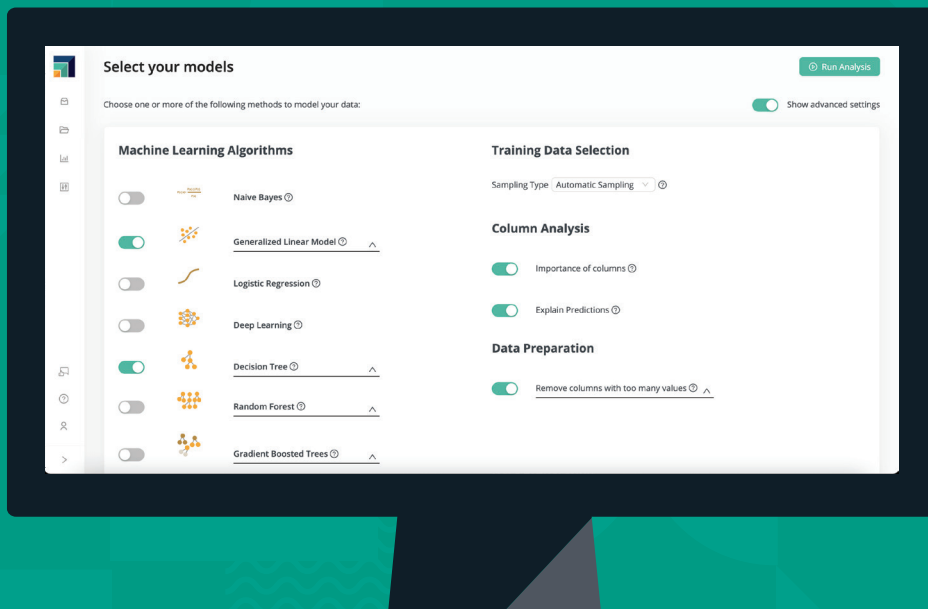### Fully integrated notebook

Leverage the power of Python and its accompanying open-source libraries to create custom solutions

Non-coder collaboration

Package your code for reuse within visual workflows or automated data science to scale your impact

Flexibility without the risk

Streamline AI governance by standardizing production and simplify auditing with a Git-based view of lineage

Select your models

Choose one or more of the following methods to model your data:

Show advanced settings

Machine Learning Algorithms

- Naive Bayes
- Generalized Linear Model
- Logistic Regression
- Deep Learning
- Decision Tree
- Random Forest
- Gradient Boosted Trees

Training Data Selection

Sampling Type: Automatic Sampling

Column Analysis

- Importance of columns
- Explain Predictions

Data Preparation

- Remove columns with too many values

Run Analysis

Search

Run

Output (2)

Decision Tree   Apply Model

Output (3)

Output

Set Role

Read Excel

Input

File Edit View Run Kernel Git Tabs Settings Help

Initial Python analysis.ipynb   tutorial.ipynb

Code   git

coding-shared-rm-base

Filter files by name

/ a18c34b4-dfe4-47d8-bd19-c12e4e7cb641 /

Name	Last Modified
AutoML	a month ago
data	a month ago
AFE_Quality...	a month ago
Data prep.r...	a month ago
Initial Pyth...	seconds ago
Model Perf...	a month ago
Model.rmm...	a month ago
Optimize q...	a month ago
Quality Pre...	a month ago
Weights.rm...	a month ago

```
os.makedirs("data", exist_ok=True)
rmpy.data.download("Settings and Quality.rmhdf5table", "./data/Settings and Quality.rmhdf5table")
Settings and Quality.rmhdf5table downloaded.
```

```
df = rmpy.data.read("data/Settings and Quality.rmhdf5table")
display(df)
```

	Quality	Temperature	S_Pressure	S_Speed	S_Rotation	Product Id
0	40.032317	49.848031	4.564443	9.694627	5.444521	<hSpy_hl.dataset.AstypeContext object at 0x7f...
1	14.991739	62.731414	7.020672	4.545744	2.668813	<hSpy_hl.dataset.AstypeContext object at 0x7f...
2	29.713303	72.263289	3.762300	2.056904	7.385801	<hSpy_hl.dataset.AstypeContext object at 0x7f...
3	28.636011	64.534637	2.766248	9.605030	2.446538	<hSpy_hl.dataset.AstypeContext object at 0x7f...
4	21.364677	81.424041	5.821376	8.992441	7.315906	<hSpy_hl.dataset.AstypeContext object at 0x7f...
9797	4.382875	59.609075	2.112033	6.254519	4.334486	<hSpy_hl.dataset.AstypeContext object at 0x7f...
9798	4.948802	61.262141	6.602571	3.594336	9.486042	<hSpy_hl.dataset.AstypeContext object at 0x7f...
9799	49.663805	66.919047	5.847640	8.462610	1.524570	<hSpy_hl.dataset.AstypeContext object at 0x7f...
9800	38.070051	77.938600	5.363073	2.497441	2.260313	<hSpy_hl.dataset.AstypeContext object at 0x7f...
9801	30.043529	70.851842	8.269202	8.783411	8.373842	<hSpy_hl.dataset.AstypeContext object at 0x7f...

9802 rows × 6 columns

```
df['Quality'].hist()
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f672167ebd0>

Simple   0   2   8   Git: idle   coding-shared-rm-base | idle   Mode: Command   Ln 1, Col 1   Initial Python analysis.ipynb

RAPIDMINER

An △ ALTAIR Company

Trusted by analytics teams from

BAIN & COMPANY

Domino's

Hewlett Packard Enterprise

LAND ROVER | JAGUAR

Lufthansa