

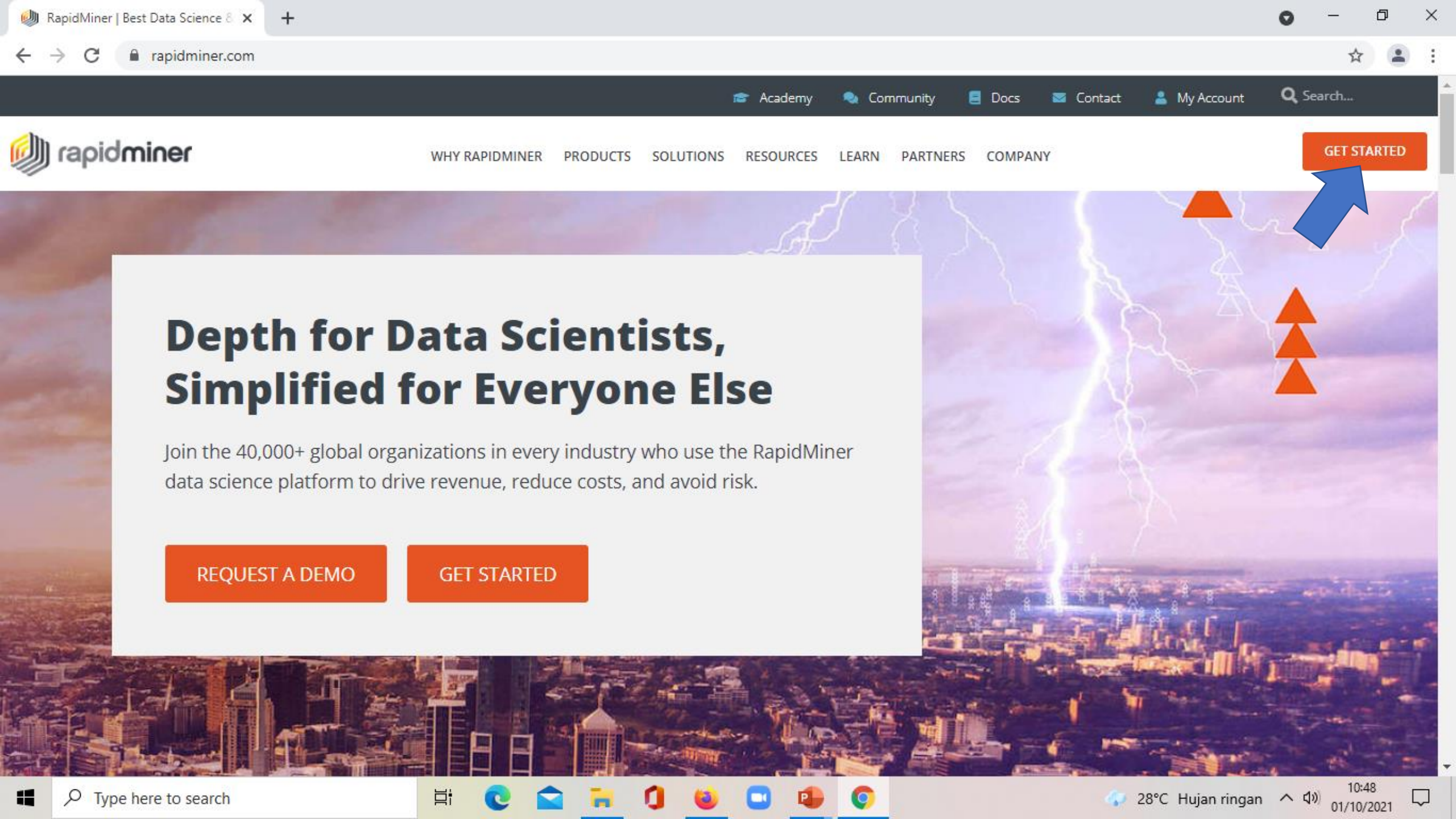


Big Data & Data Analytics

Pertemuan 2

Prodi Psikologi Kelas III & IV

Download Rapid Miner



GET STARTED

Depth for Data Scientists, Simplified for Everyone Else

Join the 40,000+ global organizations in every industry who use the RapidMiner data science platform to drive revenue, reduce costs, and avoid risk.

REQUEST A DEMO

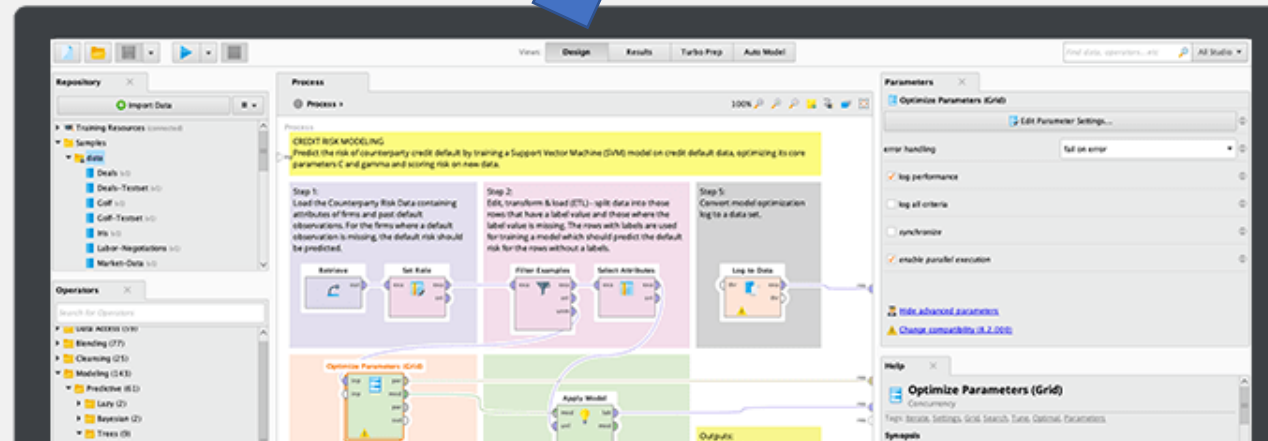
GET STARTED

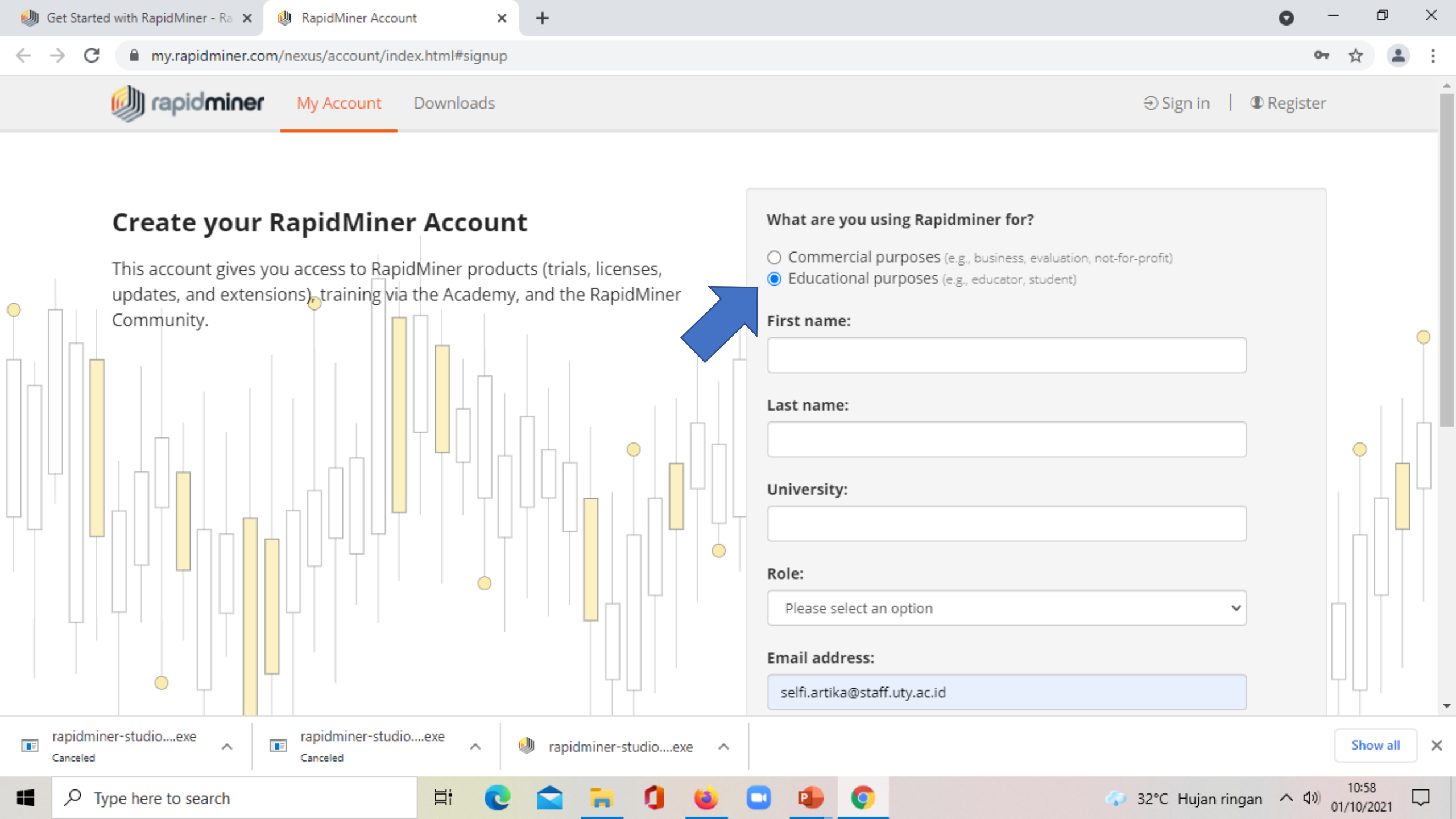


Get started with RapidMiner

RapidMiner provides all of the capabilities to support the full data science lifecycle and drive real impact across your enterprise.

DOWNLOAD





Create your RapidMiner Account

This account gives you access to RapidMiner products (trials, licenses, updates, and extensions), training via the Academy, and the RapidMiner Community.

What are you using Rapidminer for?

- ☐ Commercial purposes (e.g., business, evaluation, not-for-profit)
- ☒ Educational purposes (e.g., educator, student)

First name:

Last name:

University:

Role:

Please select an option

Email address:

selfi.artika@staff.uty.ac.id

rapidminer-studio....exe
Canceled

rapidminer-studio....exe
Canceled

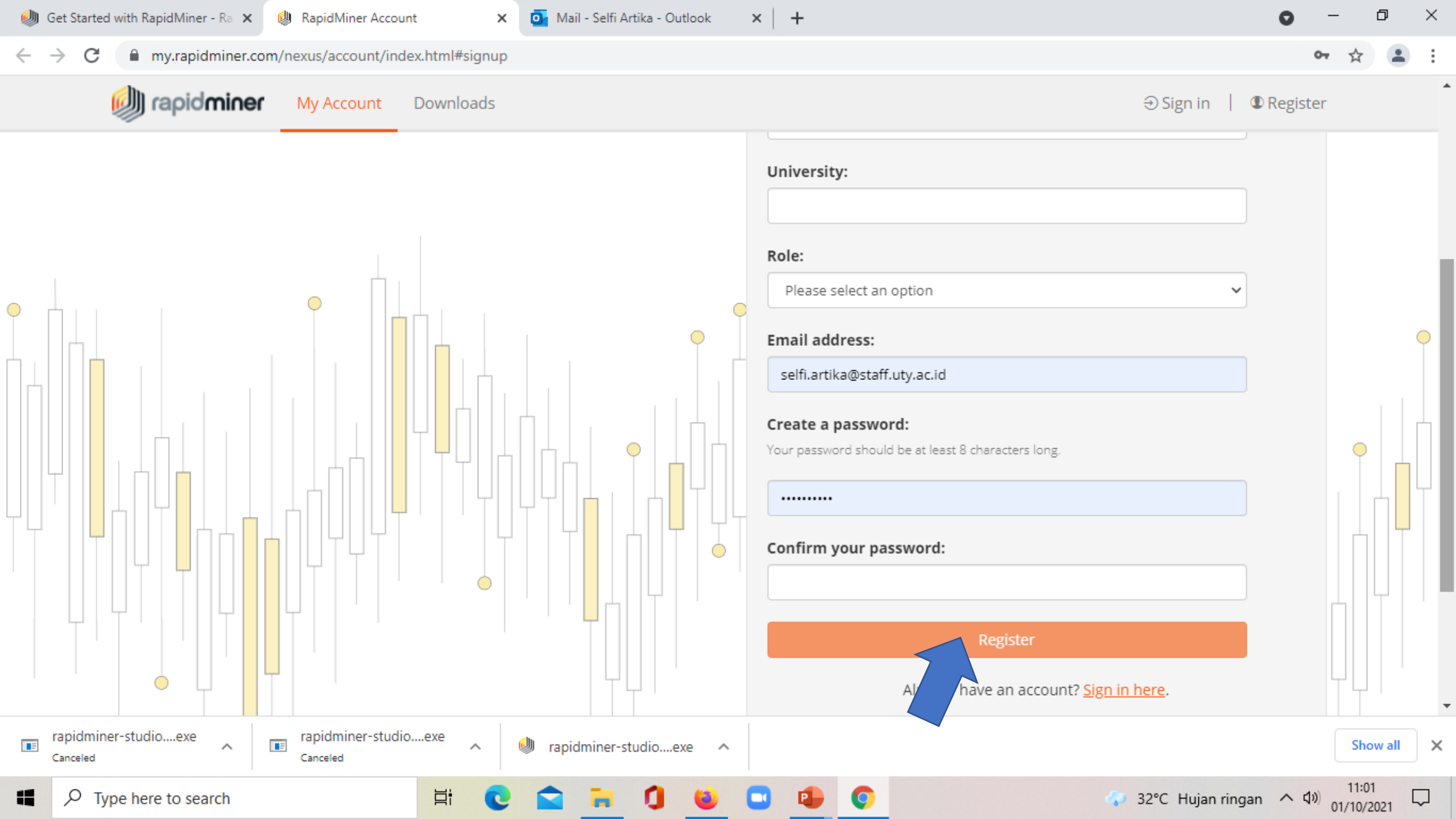
rapidminer-studio....exe

Show all

Type here to search

32°C Hujan ringan

10:58
01/10/2021



Get Started with RapidMiner - Ra

RapidMiner Account

Mail - Selfi Artika - Outlook

outlook.office.com/mail/inbox/id/AAQkADNmZmFhYmE0LTlkZDMtNDNmZi1hNGM4LTE2OTZINjdiMThhYgAQAMPFRUai4BBkCIXy%2BfAJHw%3D

UTY
UNIVERSITAS TEKNOLOGI YOGYAKARTA

Outlook

Search

Teams call

3

Selfi Artika

SA

Your browser supports setting Outlook on the Web as the default email handler. Try it now Ask again later Don't show again

New message

Delete

Archive

Junk

Sweep

Move to

Categorize

Snooze

Undo

Favorites

Inbox 31

Sent Items

Drafts

Deleted Items

Add favorite

Folders

Inbox 31

Drafts

Sent Items

Deleted Items

Focused Other Filter

RapidMiner

Verify Your Email

Welcome to RapidMiner! Thank you for r...

6:35 PM

Yesterday

IR

Ingo from RapidMiner

predictive models in 5 clicks

Wed 3:50 AM

Hi Selfi Would you believe me if I told yo...

This week

IR

Ingo from RapidMiner

Data science starts with dat...

Mon 3:12 AM

Hi there, Everyone says that 80% of data ...

Last week

Ingo from RapidMiner

Verify Your Email

Thu 9/30/2021 6:35 PM

To: Selfi Artika

logo

Welcome to RapidMiner!

Thank you for registering! Please [confirm your email address](#) to activate your account. This activation link will work for 5 days.

Verify this email

Thanks and happy mining!

Your RapidMiner Team

rapidminer-studio....exe

rapidminer-studio....exe

rapidminer-studio....exe

Type here to search

32°C Hujan ringan


11:00

01/10/2021

Get Started with RapidMiner - Ra

RapidMiner Account

my.rapidminer.com/nexus/account/index.html#home


My AccountProfileDownloadsLicenses

Sign out

Hello, Self!

Welcome to your account portal.


Products



Downloads


Get the best RapidMiner products.

Licenses




Licenses

Manage licenses and subscriptions for your RapidMiner products.



Referrals


Invite friends to use RapidMiner and get free data rows!



Educational Program


For Students and Teachers.

Account




Profile

Update your contact and profile information.




Community

Visit the RapidMiner user community.






Academy


Learn data science and RapidMiner from experts.

MarketplaceDocumentation

ImprintPrivacy Policy



Type here to search



32°C Hujan ringan

10:50
01/10/2021

Get Started with RapidMiner - Ra

RapidMiner Account

←

→

↺


my.rapidminer.com/nexus/account/index.html#downloads

🔑

☆

👤

⋮

My AccountProfileDownloadsLicenses


🔒 Sign out

Downloads


Click on a RapidMiner product of your choice to download it.

RapidMiner Studio 9.10

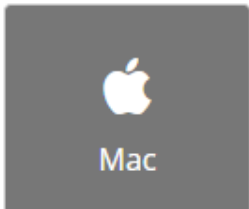
Click on your operating system to start the download:



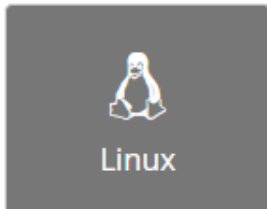
Windows
32bit



Windows
64bit



Mac




Linux
64bit
Requires: [Java 8](#)


- [Installation Guide](#)
- [Getting Started Tutorials](#)
- [Support](#)
- [Download Source](#)

RapidMiner AI Hub 9.10

Run RapidMiner AI Hub in the Cloud:



aws Amazon EC2




Microsoft Azure

- [Installation Guide](#)
- [Support](#)

Or choose your package option to start the download:

🪟

🔍 Type here to search



☁️ 32°C Hujan ringan

🔊

10:50
01/10/2021

💬



My Account

Profile

Downloads

Licenses

Sign out

Downloads

Click on a RapidMiner product of your choice to download it.

RapidMiner Studio 9.10

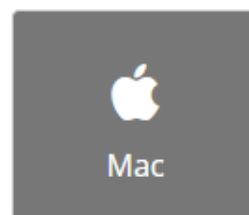
Click on your operating system to start the download:



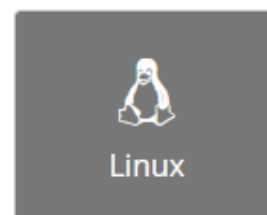
32bit



64bit



Mac



Linux

64bit

Requires: [Java 8](#)

- [Installation Guide](#)
- [Getting Started Tutorials](#)
- [Support](#)
- [Download Source](#)

RapidMiner AI Hub 9.10

Run RapidMiner AI Hub in the Cloud:



- [Installation Guide](#)


Show all



Create a RapidMiner account



You'll use your RapidMiner Account to access:


-  the Community forum
-  the Extensions Marketplace
-  free cloud storage
-  product news and updates
-  product license information

Email

Password

Confirm Password

☒ Remember my password

 **Create my Account!**

[I already have an account or license key](#)



Login to your RapidMiner Account



Login with your RapidMiner account to
download and install your licenses.

☒ Remember my password

Login and Install

[Create a new RapidMiner account](#)

[Manually enter a license key](#)

[I forgot my email or password](#)

You're all set!



Your account and license look good.
so now it's time to start using RapidMiner!

I'm ready!



Install Java

Instalasi JDK

- Double-click pada file installer JDK yang akan diinstall
- Klik "Next"



Instalasi JDK

- Klik "Next"

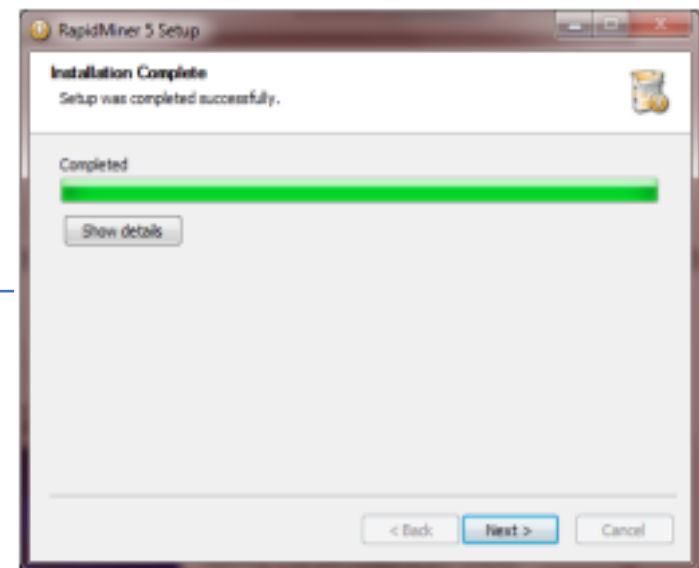
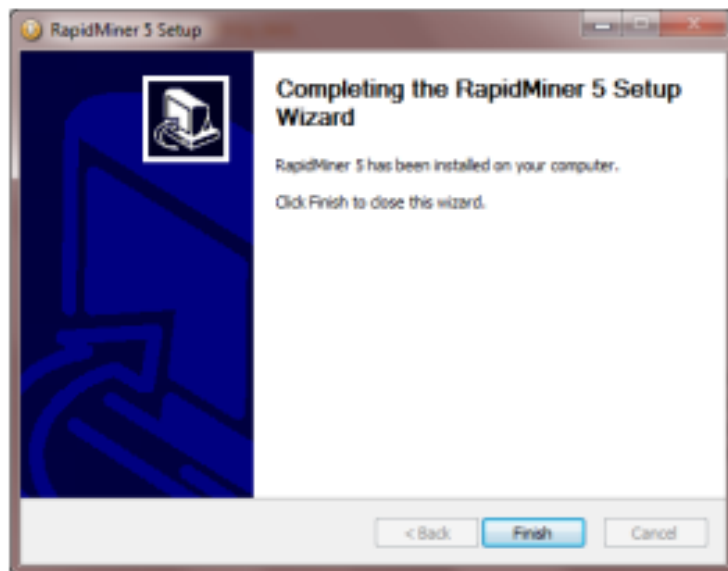
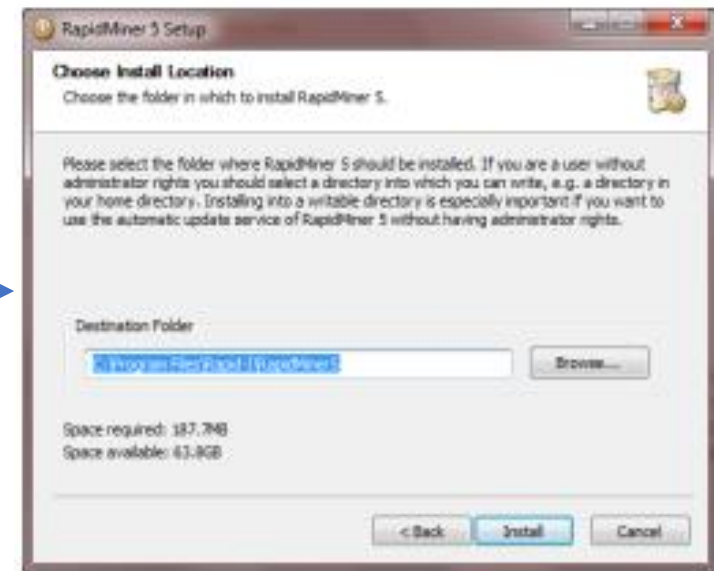
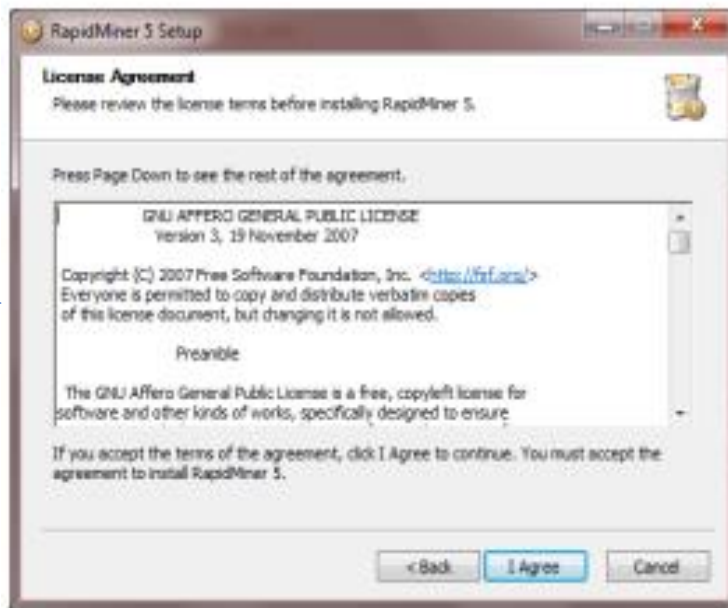


Instalasi JDK

- Instalasi JDK selesai
- Klik "Close"



Instalasi Rapidminer



Mengenal Interface Rapidminer

Setelah
Terinstall , buka
rapidminernya





Repository

+ Import Data

- Samples
- Local Repository (Local)
- DB (Legacy)

Operators

Search for Operators


- Data Access (58)
- Blending (82)
- Cleansing (28)
- Modeling (168)
- Scoring (14)
- Validation (30)
- Utility (85)
- Extensions (2)

[Get more operators from the marketplace](#)


Welcome to RapidMiner Studio!

Start Recent Learn


Start with

**Blank Process**

Start a new process from scratch in the design view.


**Turbo Prep**

Prepare your data interactively: transform, clean and combine data sets.


**Auto Model**

Build and optimize models using automated machine learning.


Choose a template to start from

**Churn Modeling**


Predict which of your customers will churn and why with a decision tree.

**Direct Marketing**


Predict response to campaigns and increase the conversion rate of your campaign.

**Credit Risk Modeling**


Model credit default risk by training an optimized Support Vector Machine (SVM) model.

**Market Basket Analysis**


Find products frequently purchased together and turn them into rules for recommendations.

**Predictive Maintenance**


Model equipment failures to schedule maintenance pre-emptively

**Price Risk Clustering**


Cluster price developments using X-Means to unveil price-risk-relationships.

**Lift Chart**


Create a lift chart to visualize the improvement that a model provides compared to guessing.


**Operationalization**


Embed predictive models into business processes to trigger the right actions automatically.

**Outlier Detection**

Detect anomalies in data resulting from a chemical analysis of wines.

**Geographic Distances**

**Medical Fraud Detection**

**Web Analytics**

etc

All Studio

init

2001

never

SYSTEM

[parameters](#)

[ibility \(9.10.000\)](#)

Studio Core

which is the outer most process.



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Samples
- Local Repository (Local)
- DB (Legacy)

Operators

Search for Operators

- Data Access (58)
- Blending (82)
- Cleansing (28)
- Modeling (168)
- Scoring (14)
- Validation (30)
- Utility (85)

[Get more operators from the Marketplace](#)

Process

Process

inp res

Your process looks empty.
Add some data first.
Drag data or operators here.

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

☒ Activate Wisdom of Crowds

Parameters

Process

logverbosity: init

logfile:

resultfile:

random seed: 2001

send mail: never

encoding: SYSTEM

[Hide advanced parameters](#)

[Change compatibility \(9.10.000\)](#)

Help

Process

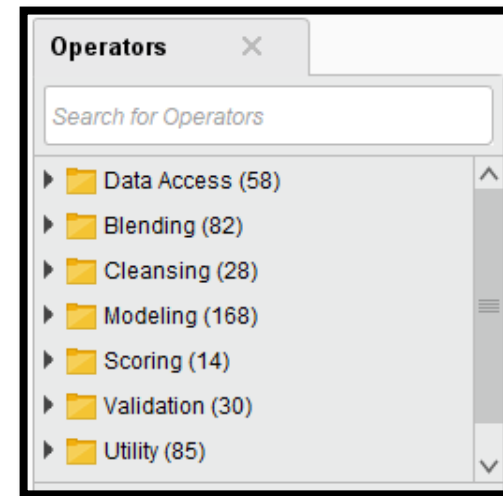
RapidMiner Studio Core

Synopsis

The root operator which is the outer most operator of every process.

Description

Operator View

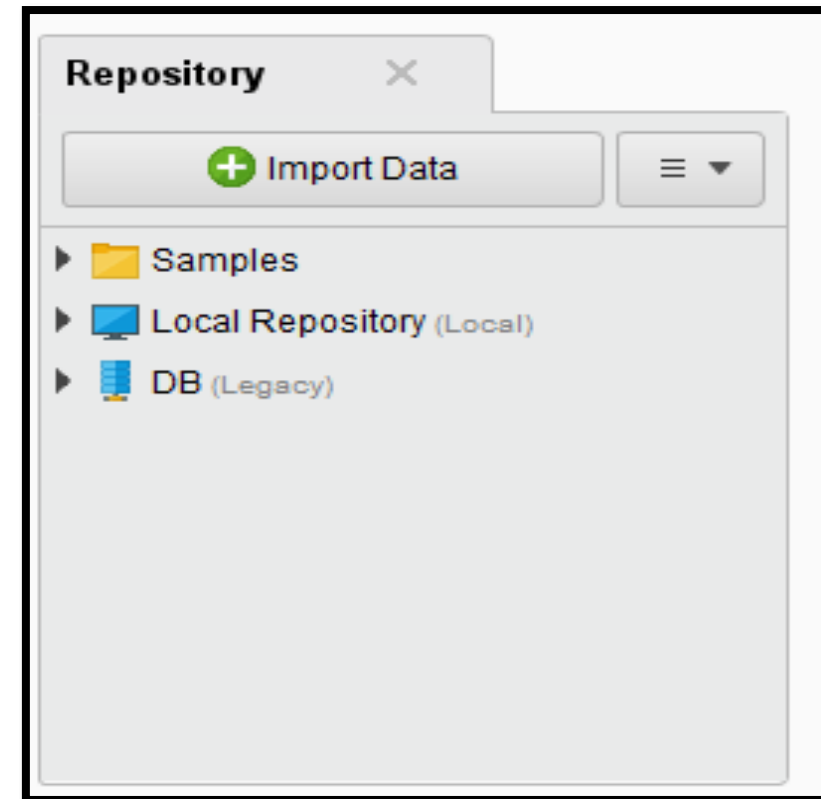


- ✖ *Process Control* : Operator ini terdiri dari operator perulangan dan percabangan yang dapat mengatur aliran proses.
- ✖ *Utility* : Operator bantuan, seperti operator macros, login, subprocesses, dan lain-lain.
- ✖ *Repository Access* : Kelompok ini terdiri dari operator-operator yang dapat digunakan untuk membaca atau menulis akses pada repository.
- ✖ *Import* : Kelompok ini terdiri dari banyak operator yang dapat digunakan untuk membaca data dan objek dari format tertentu seperti file, database, dan lain-lain.
- ✖ *Export* : Kelompok ini terdiri dari banyak operator yang dapat digunakan untuk menulis data dan objek menjadi format tertentu.

- ✖ *Data Transformation* : kelompok ini terdiri dari semua operator yang berguna untuk transformasi data dan meta data.
- ✖ *Modeling* : kelompok ini berisi proses data mining untuk menerapkan model yang dihasilkan menjadi set data yang baru.
- ✖ *Evaluation* : kelompok ini berisi operator yang dapat digunakan untuk menghitung kualitas pemodelan dan untuk data baru.

Repository View

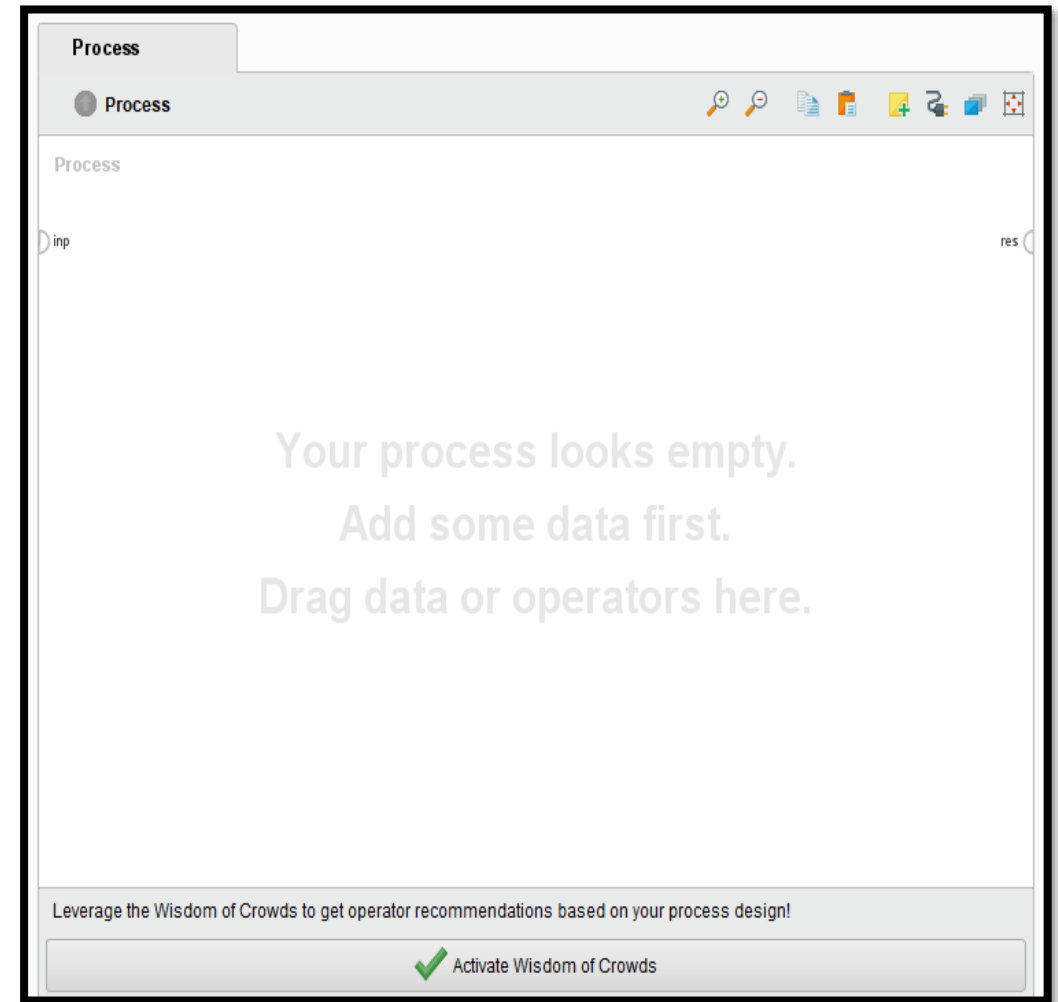
Repository View merupakan komponen utama dalam Design Perspective selain Operator View. View ini dapat Anda gunakan untuk mengelola dan menata proses Analisis Anda menjadi proyek dan pada saat yang sama juga dapat digunakan sebagai sumber data dan yang berkaitan dengan meta data.



Process View

Process View menunjukkan langkah-langkah tertentu dalam proses analisis dan sebagai penghubung langkah-langkah tersebut. Anda dapat menambahkan langkah baru dengan beberapa cara. hubungan diantara langkah-langkah ini dapat dibuat dan dilepas kembali. Pada dasarnya bekerja dengan RapidMiner ialah mendefinisikan proses analisis, yaitu dengan menunjukkan serangkaian langkah kerja tertentu. Dalam RapidMiner, komponen proses ini dinamakan sebagai operator. Operator pada RapidMiner didefinisikan sebagai berikut:

- ✖ Deskripsi dari input yang diharapkan.
- ✖ Deskripsi dari output yang disediakan.
- ✖ Tindakan yang dilakukan oleh operator pada input, yang akhirnya mengarah dengan penyediaan output.
- ✖ Sejumlah parameter yang dapat mengontrol *action performed*.



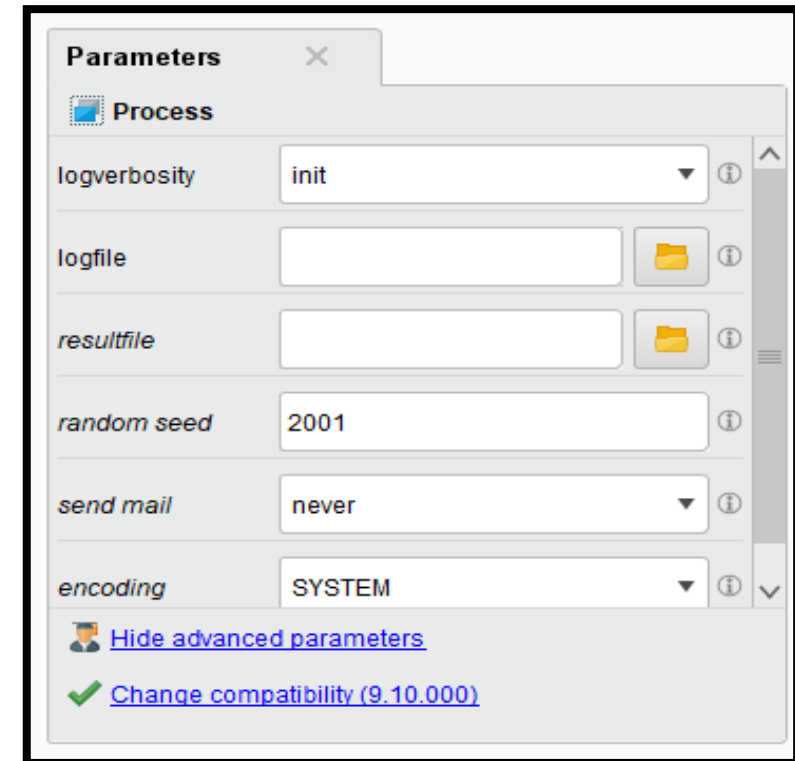
Parameters View

Beberapa operator dalam RapidMiner membutuhkan satu atau lebih parameter agar dapat diindikasikan sebagai fungsionalitas yang benar. Namun

terkadang parameter tidak mutlak dibutuhkan, meskipun eksekusi operator dapat dikendalikan dengan menunjukkan nilai parameter tertentu. Parameter view memiliki toolbar sendiri sama seperti view-view yang lain. Pada Gambar 2.12, Anda dapat melihat bahwa pada Parameter View ini terdapat beberapa ikon dan nama-nama operator terkini yang diikuti dengan aktual parameter.

Huruf tebal berarti bahwa parameter mutlak harus didefinisikan oleh analis dan tidak memiliki nilai default. Sedangkan huruf miring berarti bahwa parameter diklasifikasikan sebagai parameter ahli dan seharusnya tidak harus diubah oleh pemula untuk analisis data.

Poin pentingnya ialah beberapa parameter hanya ditunjukkan ketika parameter lain memiliki nilai tertentu.



Help & Comment View

Setiap kali Anda memilih operator pada Operator View atau Process View, maka jendela bantuan dalam Help View akan menunjukkan penjelasan mengenai operator ini. Penjelasan yang ditampilkan dalam Help View meliputi:

- ✖ Sebuah penjelasan singkat mengenai fungsi operator dalam satu atau beberapa kalimat.
- ✖ Sebuah penjelasan rinci mengenai fungsi operator.
- ✖ Daftar semua parameter termasuk deskripsi singkat dari parameter, nilai default (jika tersedia), petunjuk apakah parameter ini adalah parameter ahli serta indikasi parameter dependensi.




Sedangkan Comment View merupakan area bagi Anda untuk menuliskan komentar pada langkah-langkah proses tertentu. Untuk membuat komentar, Anda hanya perlu memilih operator dan menulis teks di atasnya dalam bidang komentar. Kemudian komentar tersebut disimpan bersama-sama dengan definisi proses Anda. Komentar ini dapat berguna untuk melacak langkah-langkah tertentu dalam rancangan nantinya.

Retrieving Data / Pengambilan Dataset

Retrieving Data Dengan Sample di Repository

Kita akan mengambil data “Titanic Training” dari Sample yang sudah ada di Repository.

- ☐ Buka folder Samples pada Repository
- ☐ Klik Folder Data
- ☐ Pilih Dataset “Titanic Training” lalu drag ke lembar Process
- ☐ Hubungkan output port pada Retrieve Titanic Training dengan port res di sebelah kanan
- ☐ Klik Run 



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

Import Data

- Labour-Regulations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training**
- Titanic Unlabeled

Operators

Search for Operators

- Data Access (58)
 - Files (20)
 - Database (3)
 - Applications (9)
 - Cloud Storage (20)
 - Retrieve
 - Store

[Get more operators from the Marketplace](#)

Process

Process

Retrieve Titanic Training

inp → [Retrieve Titanic Training] → out → res

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Parameters

Retrieve Titanic Training (Retrieve)

repository entry:

Help

Retrieve



RapidMiner Studio Core

Tags: [Load](#), [Import](#), [Read](#), [Datasets](#), [Examples](#), [Example Set](#), [Table](#), [Repository](#), [Data Access](#)

Synopsis

This Operator can access stored information in the Repository and load them into the Process.

Result History ExampleSet (Retrieve Titanic Training)

Open in  Turbo Prep  Auto Model

Filter (916 / 916 examples): all

Row No.	Survived	Age	Passenger ...	Sex	No of Sibling...	No of Parent
1	Yes	29	First	Female	0	0
2	No	2	First	Female	1	2
3	No	30	First	Male	1	2
4	No	25	First	Female	1	2
5	Yes	48	First	Male	0	0
6	Yes	63	First	Female	1	0
7	No	39	First	Male	0	0
8	Yes	18	First	Female	1	0
9	Yes	26	First	Female	0	0
10	Yes	80	First	Male	0	0
11	No	29.881	First	Male	0	0
12	No	24	First	Male	0	1
13	Yes	50	First	Female	0	1
14	Yes	32	First	Female	0	0

ExampleSet (916 examples, 1 special attribute, 6 regular attributes)

HASIL

+ Import Data

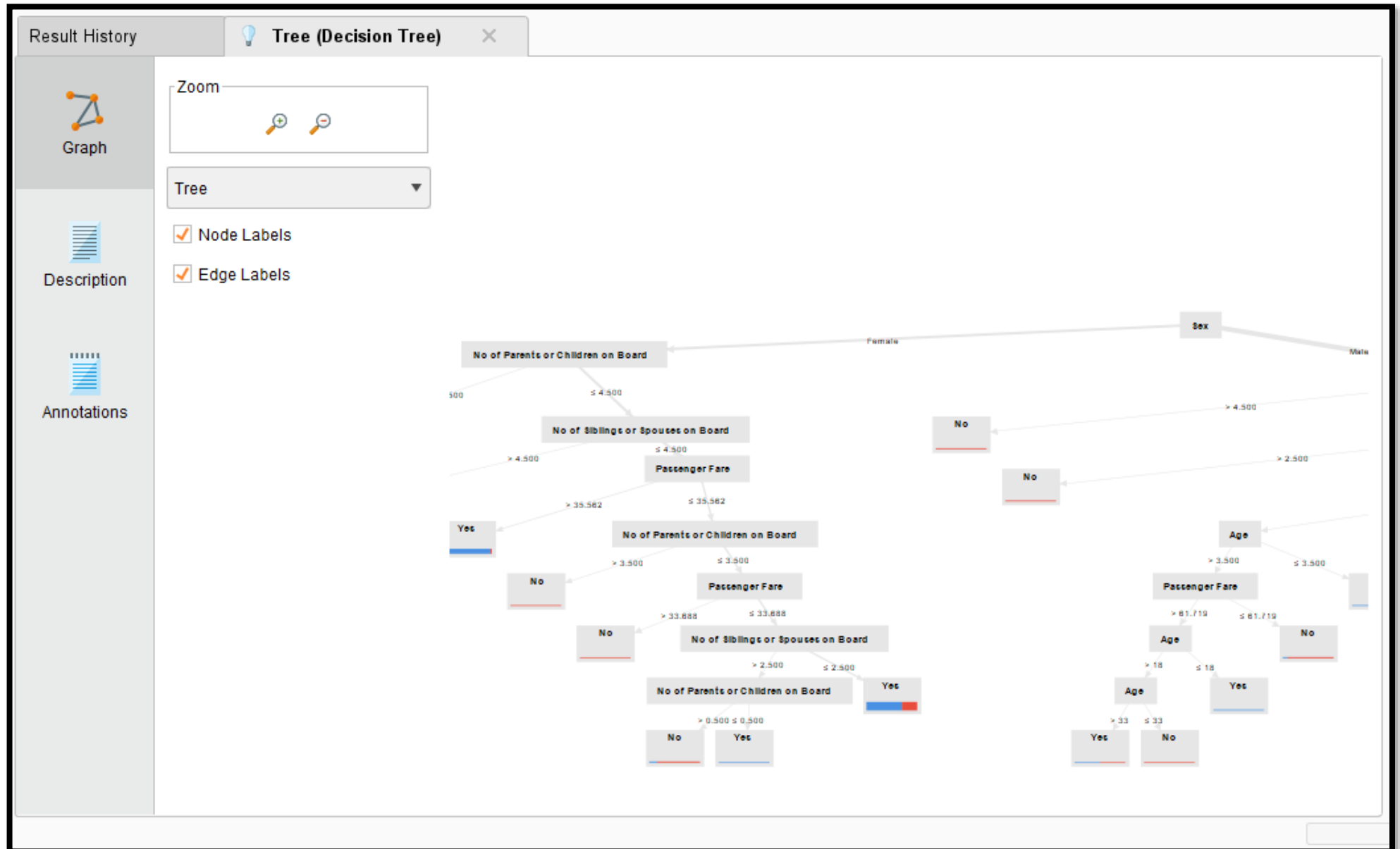
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting
- processes
- Templates
- Time Series
- Tutorials
- Local Repository (Local)
- DB (Legacy)

Membuat Pohon Keputusan

Dari data “Titanic Training” tersebut, sebanyak 349 orang dapat bertahan. Buatlah sebuah proses yang menunjukkan bahwa penumpang yang selamat itu memiliki benda – benda tertentu.

- ☐ Setelah melakukan Retrieving Data, Pilih Modelling – Predictive – Tree – Decision Tree
- ☐ Drag operator Tree ke halaman process
- ☐ Hubungkan port output pada dataset Titanic Training ke port operator process,
- ☐ Hubungkan dengan result (res) , lalu klik button play.

Hasil



Import Data

- Akan dilakukan import dataset ke rapidminer
- Dalam kuliah ini, kita akan coba gunakan dataset “status gizi” yang sudah diupload di e-learning
- **Langkah :**
 - ❑ Klik Menu Import – Pilih data yang akan diimport
 - ❑ Pilih Next – muncul data yang akan diimport
 - ❑ Pilih Next
 - ❑ Simpan Data tersebut di Local Repository – Beri Nama Filenya
 - ❑ Klik Finish



Repository

+ Import Data

- Training Resources (connected)
- Community Examples (connected)
- Sample Data
- Local Repository (Local)
- DB (Legacy)

Operators

Search for Operators

- Data Access (58)
- Blending (82)
- Cleansing (28)
- Modeling (168)
- Scoring (14)
- Validation (30)
- Utility (85)

[Get more operators from the Marketplace](#)

Import Data - Select the data location.

Select the data location.

New folder

Bookmarks	File Name	Size	Type	Last Modified
★ -- Last Directory	Status Gizi.xlsx	10 KB	Microsoft Excel Wor...	Oct 9, 2021

Status Gizi.xlsx

All Files

The selected file will be imported as: Excel [Change](#)

← Previous

Next →

Cancel

✓ Activate Wisdom of Crowds

Parameters

Process

logverbosity: init

logfile: [Folder Icon]

resultfile: [Folder Icon]

random seed: 2001

send mail: never

encoding: SYSTEM

[Hide advanced parameters](#)

[Change compatibility \(9.10.000\)](#)

Help

Process

RapidMiner Studio Core

Synopsis

The root operator which is the outer most operator of every process.

Description

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Lembar1 Cell range: A:F Select All ☒ Define header row: 1

	A	B	C	D	E	F
1	NO	PROVINSI	STATUS GIZI			
2			GIZI BURUK	GIZI KURANG	GIZI BAIK	GIZI LEBIH
3	1.000	Aceh	2.600	14.100	82.200	1.100
4	2.000	Sumatera Utara	3.100	10.100	85.200	1.700
5	3.000	Sumatera Barat	2.100	13.900	83.300	0.700
6	4.000	Riau	2.000	15.000	81.400	1.600
7	5.000	Jambi	2.400	13.200	82.600	1.900
8	6.000	Sumatera Selatan	1.900	9.300	87.200	1.600
9	7.000	Bengkulu	1.300	7.400	89.500	1.800
10	8.000	Lampung	1.600	12.400	84.500	1.600
11	9.000	Kepulauan Bang...	2.000	11.200	83.500	3.300
12	10.000	Kepulauan Riau	3.700	14.000	80.100	2.100
13	11.000	DKI Jakarta	3.000	11.300	81.300	4.400
14	12.000	Jawa Barat	2.400	12.100	84.200	1.300

Previous Next Cancel



Views

Design

Results

Turbo Prep

Auto Model

Deployments

Import Data - Format your columns.

×

Format your columns.

☐

Replace errors with missing values ⓘ

	NO <i>integer</i>	PROVINSI <i>polynomial</i>	STATUS GIZI <i>polynomial</i>	D <i>polynomial</i>	E <i>polynomial</i>	F <i>polynomial</i>
1	?	?	GIZI BURUK	GIZI KURANG	GIZI BAIK	GIZI LEBIH
2	1	Aceh	2.6	14.1	82.2	1.100000000000...
3	2	Sumatera Utara	3.1	10.1	85.2	1.7
4	3	Sumatera Barat	2.1	13.9	83.3	0.7
5	4	Riau	2	15	81.400000000000...	1.6
6	5	Jambi	2.4	13.2	82.6	1.9
7	6	Sumatera Selatan	1.9	9.300000000000...	87.2	1.6
8	7	Bengkulu	1.3	7.4	89.5	1.8
9	8	Lampung	1.6	12.4	84.5	1.6
10	9	Kepulauan Bang...	2	11.2	83.5	3.3
11	10	Kepulauan Riau	3.7	14	80.099999999999...	2.1
12	11	DKI Jakarta	3	11.3	81.3	4.400000000000...

✓

no problems.

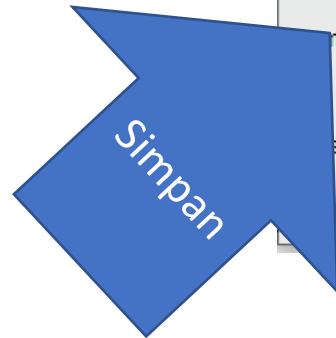
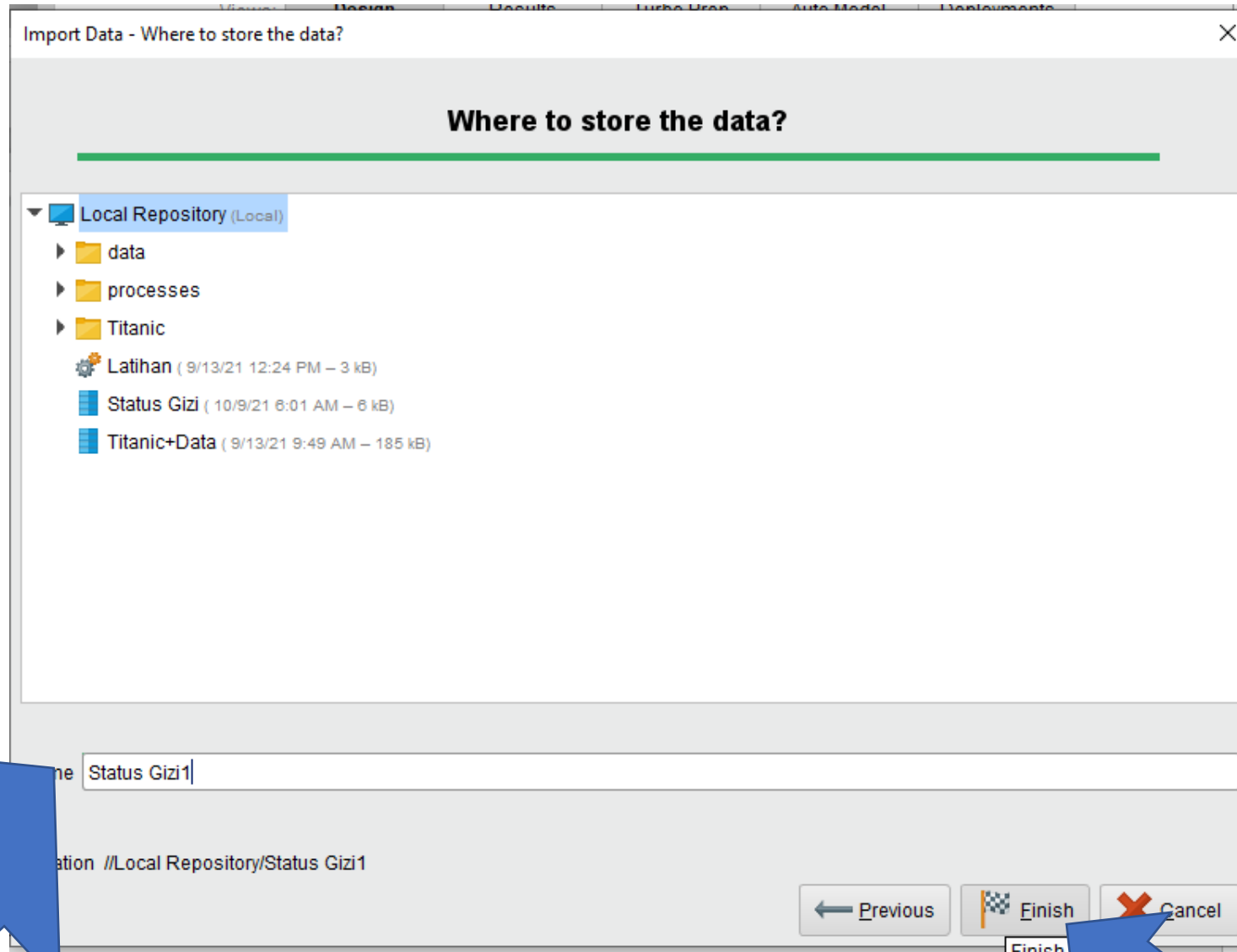
← Previous

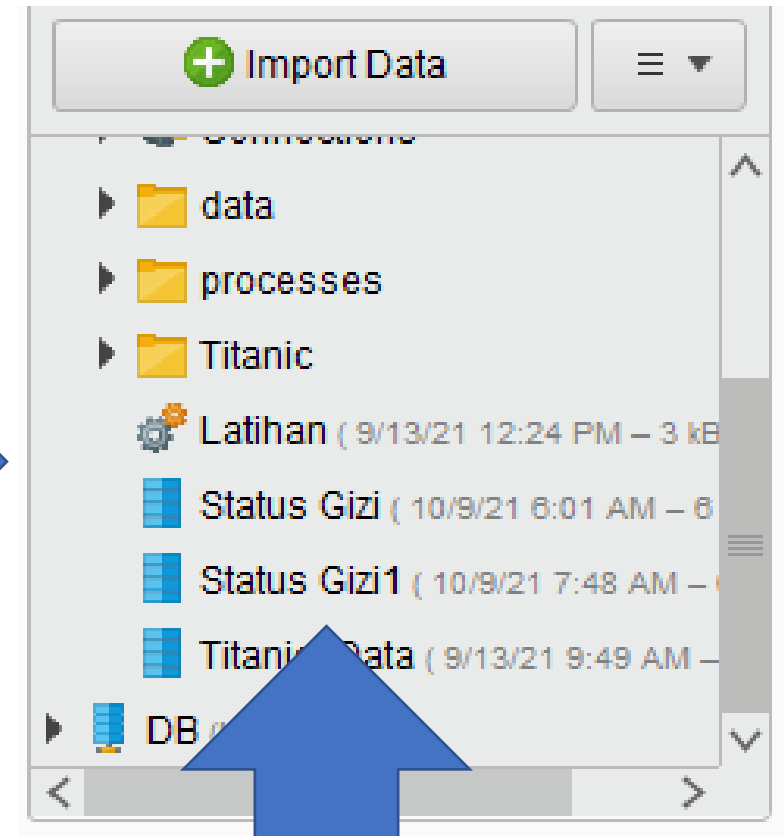
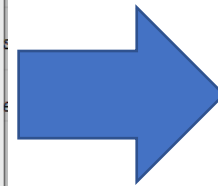
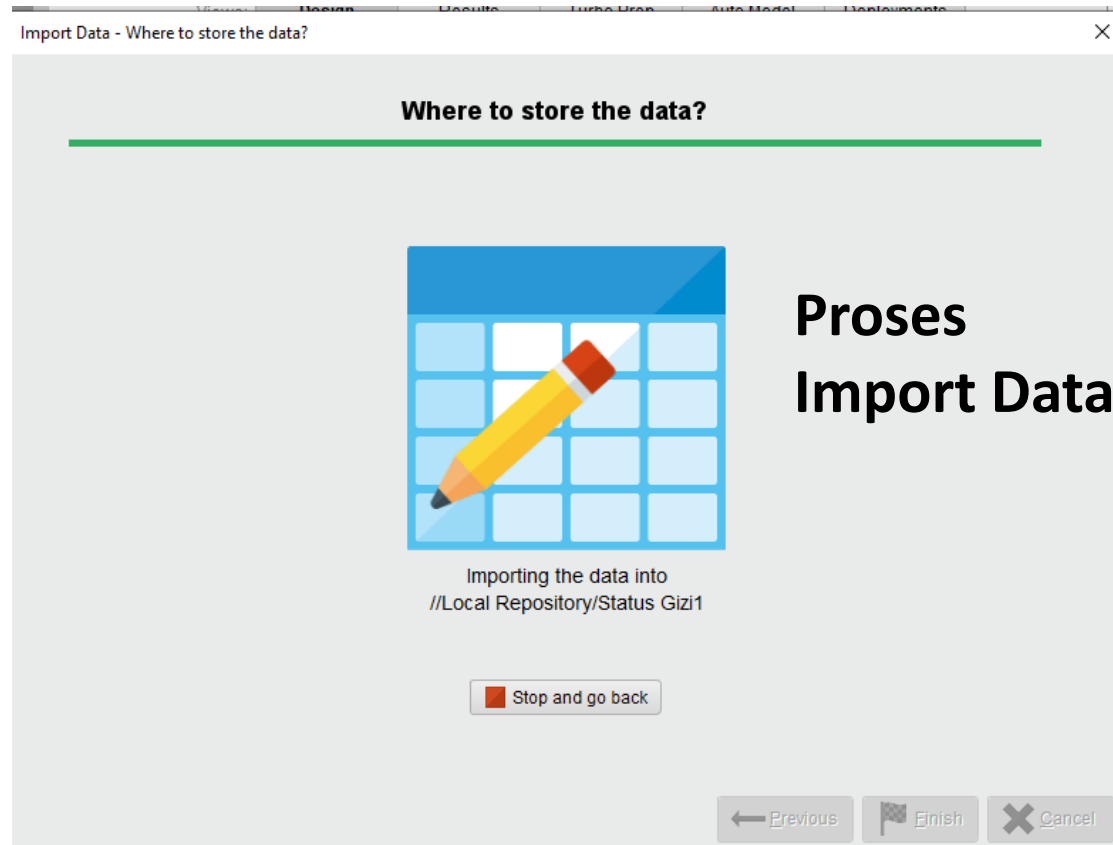
Next →

✗

Cancel

5





Data Telah Selesai Di Import

