

Structure

- 1. Data preparation
 - 1.1. Import
 - 1.2. Filter
 - 1.3. Enrich
 - 1.4. Random test and train split
 - 1.5. k-fold cross-validation
- 2. Data prediction
 - 2.1 Import Libraries and Data
 - 2.2 Regression models and hyperparameter grids
 - 2.3 Cross-validation with GridSearchCV
 - 2.4 Model performance evaluation
 - 2.5 Select best model
 - 2.6 Predict future values
- 3. Data prescription
 - 3.1 Baseline Scheduling
 - 3.2 Smart Stochastic scheduling (SSS)
 - 3.2.1 Loss function
 - 3.2.2 Stochastic model
 - 3.2.3 Stochastic analysis
 - 3.5 GUI

1. Data preparation

1.1. Import

Load the needed data:

- job_logs.csv = Logs of historical Jobs with Timestamps
- weekly_releases.csv = List of Jobs that are executed every week

```
In [1]: import pandas as pd

# Load historical data into a Pandas DataFrame
logs = pd.read_csv('data/jobs_logs.csv')

# Load scheduled job data into a Pandas DataFrame
releases = pd.read_csv('data/weekly_releases.csv')
```

1.2 Filter

Filter everything out that is not needed

```
In [2]: # Keep only the rows in logs whose names appear in releases.
logs = logs[logs['ProcessName'].isin(releases['Name'])]

# Remove outliers
```

```
In [3]: # Calculate the Interquartile range (IQR)
Q1 = logs['DateTime'].quantile(0.25)
Q3 = logs['DateTime'].quantile(0.75)
IQR = Q3 - Q1

# Remove outliers
filtered_logs = logs[~((logs['DateTime'] < (Q1 - 1.5 * IQR)) | (logs['DateTime'] > (Q3 + 1.5 * IQR)))]
```

1.3 Enrich

Adding min, max, mean and median execution time as well as the weekday and calendarweek to the data

```
In [4]: # Format columns as 'datetime'
filtered_logs.loc[:, 'date'] = pd.to_datetime(filtered_logs['StartTime']).dt.date.copy()
filtered_logs.loc[:, 'time'] = pd.to_datetime(filtered_logs['StartTime']).dt.time.copy()

# Add weekday
filtered_logs.loc[:, 'weekday'] = pd.to_datetime(filtered_logs['StartTime']).dt.dayofweek.copy()

# Add calendarweek
filtered_logs.loc[:, 'calendarweek'] = pd.to_datetime(filtered_logs['StartTime']).dt.isocalendar().week.astype(int)

# Rename the 'DateTime' column to 'ExecTime'
filtered_logs = filtered_logs.rename(columns={'DateTime': 'ExecTime'})

# Sort DataFrame by process name and date
filtered_logs = filtered_logs.sort_values(by=['ProcessName', 'date'])

# Add Last_ExecTime
filtered_logs.loc[:, 'last_ExecTime'] = filtered_logs.groupby('ProcessName')['ExecTime'].shift(1).copy()

# Add mean_ExecTime
filtered_logs.loc[:, 'mean_ExecTime'] = filtered_logs.groupby('ProcessName')['ExecTime'].transform('mean').copy()

# Add max_ExecTime
filtered_logs.loc[:, 'max_ExecTime'] = filtered_logs.groupby('ProcessName')['ExecTime'].transform('max').copy()

# Add min_ExecTime
filtered_logs.loc[:, 'min_ExecTime'] = filtered_logs.groupby('ProcessName')['ExecTime'].transform('min').copy()

# Add median_ExecTime
filtered_logs.loc[:, 'median_ExecTime'] = filtered_logs.groupby('ProcessName')['ExecTime'].transform('median').copy()

# Replace NaN with zero on all columns
filtered_logs = filtered_logs.fillna(0)

# filtered_logs.head()
```

```
C:\Users\ryadmi\AppData\Local\Temp\2\ipykernel_4568\1741696983.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using loc[row_index,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
filtered_logs.loc[:, 'date'] = pd.to_datetime(filtered_logs['StartTime']).dt.date.copy()
C:\Users\ryadmi\AppData\Local\Temp\2\ipykernel_4568\1741696983.py:9: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using loc[row_index,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
filtered_logs.loc[:, 'time'] = pd.to_datetime(filtered_logs['StartTime']).dt.time.copy()
C:\Users\ryadmi\AppData\Local\Temp\2\ipykernel_4568\1741696983.py:9: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using loc[row_index,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
filtered_logs.loc[:, 'weekday'] = pd.to_datetime(filtered_logs['StartTime']).dt.dayofweek.copy()
C:\Users\ryadmi\AppData\Local\Temp\2\ipykernel_4568\1741696983.py:9: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using loc[row_index,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
filtered_logs.loc[:, 'calendarweek'] = pd.to_datetime(filtered_logs['StartTime']).dt.isocalendar().week.astype(int).copy()
```

```
In [5]: # Save filtered_logs to csv('filtered_logs.csv', sep='t')
```

```
In [6]: from sklearn.preprocessing import LabelEncoder

# Preprocess the data
filtered_logs['StartTime'] = pd.to_datetime(filtered_logs['StartTime'])
filtered_logs['weekday'] = pd.to_datetime(filtered_logs['date'])
filtered_logs['calendarweek'] = filtered_logs['StartTime'].dt.weekday
filtered_logs['calendarweek'] = filtered_logs['StartTime'].dt.isocalendar().week
filtered_logs['calendarweek'] = pd.to_datetime(filtered_logs['StartTime']).dt.isocalendar().week.astype(int)

# Label encoding for categorical features
encoder = LabelEncoder()
filtered_logs['ProcessName'] = encoder.fit_transform(filtered_logs['ProcessName'])

# Define the features and target
X = filtered_logs[['ProcessName', 'weekday', 'calendarweek', 'last_ExecTime', 'mean_ExecTime', 'max_ExecTime', 'min_ExecTime', 'median_ExecTime']]
y = filtered_logs['ExecTime']
```

1.4 Random test and train split

The idea behind a test and train split is to evaluate the performance of a machine learning model on unseen data. The aim of machine learning is to learn patterns and relationships from a dataset that can be used to make predictions on new, unseen data. To ensure that the model can generalize well to new data, it is important to test its performance on a separate set of data that was not used for training.

Source: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html

```
In [7]: from sklearn.model_selection import train_test_split, GridSearchCV, KFold, cross_val_score
from sklearn.preprocessing import StandardScaler

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Standardize the data
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)
```

1.5 k-fold cross-validation

K-fold cross-validation is a model evaluation technique that helps assess a model's performance on unseen data. It is widely used to reduce the risk of overfitting, estimate the generalization error, and select the best model. The dataset is divided into k equally-sized folds (or partitions), where k is a positive integer (in our case five). The model is trained on k-1 folds and validated on the remaining fold. This process is repeated k times, ensuring that each fold is used as the validation set exactly once.

In scikit-learn's KFold class, the parameters n_splits, shuffle, and random_state have the following meanings:

- n_splits:** The number of folds (k) to divide the dataset into. This parameter should be set to a positive integer (typically between 5 and 10). Higher values of k result in a lower bias but higher variance in model evaluation, whereas lower values of k can lead to a higher bias but lower variance.
- shuffle:** This is a boolean parameter. When set to True, the dataset will be shuffled before splitting into folds. Shuffling is useful when the dataset has an inherent order that might affect the model's performance during cross-validation. When set to False, the dataset will not be shuffled before splitting, and the folds will be created by sequentially selecting data points.
- random_state:** This parameter is used to control the randomness when shuffling the dataset. If set to an integer, it serves as the seed for the random number generator, ensuring that the shuffling is consistent across multiple runs. This helps in achieving reproducible results. When set to None, the random number generator uses a different seed in each run, resulting in a different shuffling of the dataset.

Source: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html

```
In [8]: # Set up K-fold cross-validation
kf = KFold(n_splits=5, shuffle=True, random_state=42)
```

2. Data prediction

2.1 Import Libraries and Data

- Scikit-learn (sklearn):** Scikit-learn is a popular machine learning library for Python. It provides a comprehensive collection of machine learning algorithms for classification, regression, clustering, dimensionality reduction, and other tasks. Scikit-learn also includes tools for preprocessing data, model selection, and evaluation, making it a complete package for building and deploying machine learning models. It is built on top of Numpy, Scipy, and Matplotlib, and it is designed to be easy to use and highly efficient. Source: <https://scikit-learn.org/stable/install.html>
- XGBoost:** XGBoost (eXtreme Gradient Boosting) is an open-source, highly efficient, and flexible machine learning library designed for gradient boosting trees. It was developed by Tianqi Chen and Carlos Guestrin and has gained significant popularity in the machine learning community due to its performance and scalability. Source: https://xgboost.ai/docs/en/stable/python/python_intro.html

```
In [9]: from sklearn.linear_model import LinearRegression, Ridge, Lasso, ElasticNet
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor, GradientBoostingRegressor
from sklearn.svm import SVC
from sklearn.metrics import mean_squared_error, r2_score
from sklearn.model_selection import train_test_split, GridSearchCV, KFold, cross_val_score
from sklearn.preprocessing import StandardScaler, LabelEncoder
import xgboost as xgb
import time
```

2.2 Regression models and hyperparameter grids

Regression models: Regression is a supervised learning technique used for predicting continuous numerical values. It is a statistical method that models the relationship between a dependent variable (such as a constant value as the target or response variable) and one or more independent variables (also known as the predictors or explanatory variables). Source: https://scikit-learn.org/stable/supervised_learning.html

Here are brief descriptions of some popular regression models:

- Linear Regression:** Linear regression is a simple and widely used linear approach to modeling the relationship between a dependent variable and one or more independent variables. It assumes a linear relationship between the variables and tries to fit a straight line through the data that minimizes the sum of the squared errors.
- Ridge Regression:** Ridge regression is a regularized version of linear regression that adds a penalty term to the cost function to prevent overfitting. It uses L2 regularization to shrink the coefficients towards zero, which reduces the variance of the estimates and improves the model's generalization performance.
- Lasso Regression:** Lasso regression is another regularized version of linear regression that uses L1 regularization to shrink the coefficients towards zero. It can be used for feature selection, as it tends to drive some coefficients to exactly zero, which removes the corresponding predictors from the model.
- Elastic Regression:** Elastic regression is a combination of ridge and lasso regression that uses a mixture of L1 and L2 regularization to balance the benefits of both. It can be tuned to favor either L1 or L2 regularization, or a mixture of both.
- Decision Tree Regression:** Decision tree regression is a non-parametric method that recursively splits the data into subsets based on the values of the predictors, and fits a simple model (such as a constant value) to each subset. It is easy to interpret and can capture non-linear relationships between the variables, but can suffer from overfitting.
- Random Forest Regression:** Random forest regression is an ensemble method that combines multiple decision tree regressors to improve the prediction accuracy and reduce overfitting. It randomly selects subsets of the data and variables to build each tree, and aggregates the predictions of the trees to produce the final output.
- Gradient Boosting Regression:** Gradient boosting regression is another ensemble method that combines multiple weak models (such as decision trees) to produce a strong predictive model. It trains each model sequentially to correct the errors of the previous model, and uses gradient descent to minimize the loss function.
- XGBoost:** XGBoost is an optimized implementation of gradient boosting that uses a variety of techniques (such as regularization, early stopping, parallel processing) to improve the performance and scalability of the algorithm. It is often used in machine learning competitions and has been shown to achieve state-of-the-art results on a wide range of datasets.

Hyperparameter tuning: Hyperparameter tuning is the process of finding the best set of hyperparameters for a machine learning model to optimize its performance. Hyperparameters are configuration variables that govern the training process and cannot be learned directly from the data. They can have a significant impact on the model's performance, and thus, finding the right combination is crucial for building a successful model. In this case the hyperparameter tuning is done with GridSearchCV. Source: https://scikit-learn.org/stable/modules/generated/sklearn.grid_search.html

```
In [10]: # every regression model with its hyperparameter grids
models = [
    ('name': 'Linear', 'model': LinearRegression(), 'param_grid': {}), # no hyperparameters
    ('name': 'Ridge', 'model': Ridge(), 'param_grid': {'alpha': [0.1, 1, 10]}),
    ('name': 'Lasso', 'model': Lasso(), 'param_grid': {'alpha': [0.1, 1, 10]}),
    ('name': 'Elastic Net', 'model': ElasticNet(), 'param_grid': {'alpha': [0.1, 1, 10], 'l1_ratio': [0.1, 0.5, 1]}),
    ('name': 'Decision Tree', 'model': DecisionTreeRegressor(), 'param_grid': {'max_depth': [None, 5, 10]}),
    ('name': 'Random Forest', 'model': RandomForestRegressor(), 'param_grid': {'n_estimators': [10, 50, 100]}),
    ('name': 'Gradient Boosting', 'model': GradientBoostingRegressor(), 'param_grid': {'n_estimators': [50, 100, 200]}),
    ('name': 'XGBoost', 'model': xgb.XGBRegressor(objective='reg:squarederror'), 'param_grid': {
        'n_estimators': [50, 100, 150],
        'learning_rate': [0.01, 0.1, 0.2],
        'max_depth': [3, 5, 7],
        'subsample': [0.8, 1.0]
    })
]
```

```
In [11]: # create a dataframe to save results
results = []
'Model': [],
'Mean Squared Error': [],
'R2 Score': [],
results_df = pd.DataFrame(results)
```

2.3 Cross-validation with GridSearchCV

We train and evaluate multiple models using cross-validation and grid search for hyperparameter tuning.

Source: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html#sklearn.model_selection.GridSearchCV

```
In [12]: models_dict = {}

for i in models:
    print(f"Training and evaluating {i['name']}...")
    grid_search = GridSearchCV(model=i['model'], n_param_grid=i['param_grid'], scoring='neg_mean_squared_error', cv=kfold, verbose=1)
    grid_result = grid_search.fit(X_train, y_train)
    best_model = grid_result.best_estimator_
    y_pred = best_model.predict(X_test)
```

```
# Calculate evaluation metrics
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
```

```
# Measure the time it takes to train the model
end_time = time.time()
learning_time = end_time - start_time
```

```
# Assign values
key = f'{i["name"]}'
value = best_model
models_dict[key] = value
```

```
print(f"{i['name']}")
print(f"Mean Squared Error: {mse:2f}")
print(f"R2 Score: {r2:2f}")
print(f"Learning Time: {learning_time:2f}")
print(f"{i['name']}")
```

```
# Create a new DataFrame with the results for the model
model_results = pd.DataFrame({'Model': key, 'Mean Squared Error': [mse], 'R2 Score': [r2], 'Learning Time': [learning_time]})
```

```
# Concatenate the new DataFrame to the existing results DataFrame
results_df = pd.concat([results_df, model_results], ignore_index=True)
```

print(results_df)

Training and evaluating Linear...

Linear:
Mean Squared Error: 296.06
R2 Score: 0.73
Learning Time: 0.49s

Training and evaluating Ridge...

Ridge:
Mean Squared Error: 296.06
R2 Score: 0.73
Learning Time: 0.78s

Training and evaluating Lasso...

Lasso:
Mean Squared Error: 296.06
R2 Score: 0.73
Learning Time: 1.03s

Training and evaluating Elastic Net...

Elastic Net:
Mean Squared Error: 296.10
R2 Score: 0.73
Learning Time: 2.63s

Training and evaluating Decision Tree...

Decision Tree:
Mean Squared Error: 244.71
R2 Score: 0.77
Learning Time: 6.82s

Training and evaluating Random Forest...

Random Forest:
Mean Squared Error: 237.34
R2 Score: 0.78
Learning Time: 786.63s

Training and evaluating Gradient Boosting...

Gradient Boosting:
Mean Squared Error: 235.58
R2 Score: 0.78
Learning Time: 3307.55s

Training and evaluating XGBoost...

XGBoost:
Mean Squared Error: 235.17
R2 Score: 0.78
Learning Time: 1263.44s

	Model	Mean Squared Error	R2 Score	Learning Time
0	Linear	296.063872	0.72279	0.489583
1	Ridge	296.063826	0.72279	0.780494
2	Lasso	296.064447	0.72279	1.034223
3	Elastic Net	296.102621	0.72244	2.629388
4	Decision Tree	244.713156	0.774581	6.819646
5	Random Forest	237.342286	0.781371	786.628118
6	Gradient Boosting	235.579804	0.782095	3307.550284
7	XGBoost	235.165783	0.783376	1263.440491

2.4 Model performance evaluation

Visualize the results from the GridSearch.

```
In [13]: import matplotlib.pyplot as plt
import seaborn as sns

# Set plot style
sns.set(style='whitegrid')

# Create subplots
fig, ax = plt.subplots(1, 2, figsize=(8, 3))

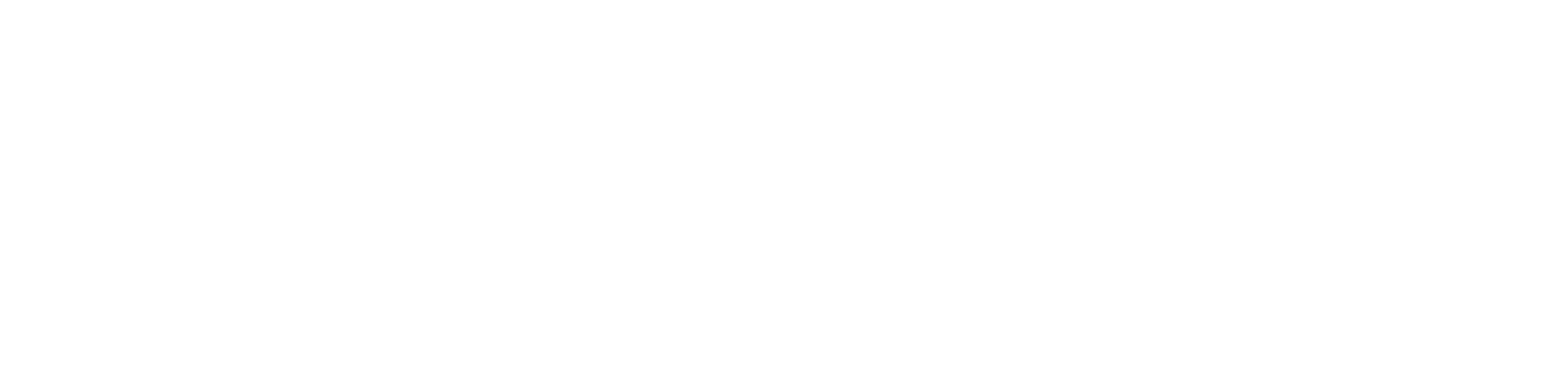
# Plot Mean Squared Error
sns.barplot(x='Model', y='Model', data=results_df, ax=ax[0], palette='Blues_d')
ax[0].set_xlabel('Mean Squared Error')
ax[0].set_ylabel('Regression Model')
ax[0].set_title('Mean Squared Error by Model')

# Plot R2 Score
sns.barplot(x='R2 Score', y='Model', data=results_df, ax=ax[1], palette='Greens_d')
ax[1].set_xlabel('R2 Score')
ax[1].set_ylabel('Regression Model')
ax[1].set_title('R2 Score by Model')
ax[1].set_xlabel('R2 Score')
ax[1].set_ylabel('Regression Model')
ax[1].set_title('R2 Score by Model')

# Adjust layout
plt.tight_layout()

# Save the plot as an SVG file
plt.savefig('images/model_evaluation.svg', format='svg')

# Display the plot
plt.show()
```



2.5 Select best model

```
In [14]: # Create a new column with the product of 'Mean Squared Error' and 'R2 Score'
results_df['R2 x 1/MSE'] = results_df['R2 Score'] * 1 / results_df['Mean Squared Error']

# Find the index of the row with the highest product value
idx_max_mse_r2 = results_df['R2 x 1/MSE'].idxmax()

# Get the model name for the highest-scoring row
best_model_name = list(results_df.loc[idx_max_mse_r2, 'Model'])[0]
```

```
# Print the model name and the product value for the highest-scoring row
print(f"Model name: {best_model_name} with R2 x 1/MSE: {results_df.loc[idx_max_mse_r2, 'R2 x 1/MSE'].values[0]:.2f} * 1e8")
```

```
# Assign the best model
model = models_dict[best_model_name]
```

```
# Assign mse
model_mse = results_df.loc[idx_max_mse_r2, 'Mean Squared Error']
print(model, model_mse)
```

```
Model with highest R2 x 1/MSE value: XGBoost
Highest R2 x 1/MSE value: 0.80 * 1e8
XGBRegressor(base_score=None, booster=None, callbacks=None,
              colsample_bylevel=None, colsample_bytree=None,
              enable_categorical=False, eval_metric=None, feature_types=None,
              gamma=None, gpu_id=None, grow_policy=None, importance_type=None,
              interaction_constraints=None, learning_rate=0.1, max_bin=None,
              max_cat_threshold=None, max_cat_to_onehot=False,
              max_delta_step=None, max_depth=7, max_leaves=None,
              min_child_weight=None, missing=None, monotone_constraints=None,
              n_estimators=50, n_jobs=None, num_parallel_tree=None,
              predictor=None, random_state=None, ...)
```

2.6 Predict future values

```
In [15]: import datetime

# Load scheduled job data into a Pandas DataFrame
future_jobs = pd.read_csv('data/weekly_releases.csv')

# Sort by Name
future_jobs = future_jobs.sort_values(by='Name', ascending=True)

# Label encoding for categorical features
encoder = LabelEncoder()
future_jobs['Name'] = encoder.fit_transform(future_jobs['Name'])

# Add a new column with a constant value
future_jobs['StartTime'] = datetime.datetime.now()
```

```
# Group by name and select row with latest date time
df1_latest = filtered_logs.groupby('ProcessName').apply(lambda x: x.loc[x['StartTime'].idxmax()]).reset_index()
```

```
# Convert start column to datetime type
df2['start'] = pd.to_datetime(df1['start'])

# Merge dataframes
future_jobs = pd.merge(future_jobs, df1_latest[['ProcessName', 'last_ExecTime', 'mean_ExecTime', 'max_ExecTime', 'min_ExecTime', 'median_ExecTime']], on='ProcessName', how='left')
```

Create a new dataframe with the upcoming jobs.

```
In [16]: # Preprocess the future job data
future_jobs['StartTime'] = pd.to_datetime(future_jobs['StartTime'])
future_jobs['weekday'] = future_jobs['StartTime'].dt.weekday
future_jobs['calenderweek'] = future_jobs['StartTime'].dt.isocalendar().week

# Define the features
X_future = future_jobs[['ProcessName', 'weekday', 'calenderweek', 'last_ExecTime', 'mean_ExecTime', 'max_ExecTime', 'min_ExecTime', 'median_ExecTime']]

# Standardize the data
X_future = scaler.transform(X_future)
```

```
# Make a prediction for the future job
future_ExecTime_pred = model.predict(X_future)
```

```
# Add a new column with the predicted values
future_jobs['ExecTime'] = future_ExecTime_pred

# Replace NaN with zero on all columns
future_jobs = future_jobs.fillna(0)

# future_jobs.head()
```

```
In [17]: import numpy as np

# Define dataframe
future_jobs_tenth = future_jobs.copy()

# Convert duration from seconds to minutes and round to nearest whole minute
future_jobs_tenth['ExecTime'] = np.ceil(future_jobs_tenth['ExecTime'] / 60 / 10)
future_jobs_tenth['start'] = sum(x[1, 5, 1] for i in range(max_workers))
future_jobs_tenth['mean_ExecTime'] = np.ceil(future_jobs_tenth['mean_ExecTime'] / 60 / 10)
future_jobs_tenth['max_ExecTime'] = np.ceil(future_jobs_tenth['max_ExecTime'] / 60 / 10)
future_jobs_tenth['min_ExecTime'] = np.ceil(future_jobs_tenth['min_ExecTime'] / 60 / 10)
future_jobs_tenth['median_ExecTime'] = np.ceil(future_jobs_tenth['median_ExecTime'] / 60 / 10)
```

```
# Head Table
future_jobs_tenth.head()
```

```
In [18]: # Create List of tuples with (Id, ExecTime)
jobs = []
for index, row in future_jobs_tenth.iterrows():
    jobs.append((row['Id'], int(row['ExecTime'])))

amount = len(jobs)
Total = future_jobs_tenth['ExecTime'].sum()
print(Total*10, 'min Execution Time', amount, 'Tasks')
2500.0 min Execution Time, 250 Tasks
```

3. Data prescription

Problem definition:

Given:

- List of jobs with execution time in seconds
- A week has 7 days. A day have 24 hours
- Each worker can work 24/7

Obj:

Minimize the amount of workers

Constr:

- Each job has to be executed once
- A worker can only execute one job simultaneously

Output:

A schedule for each worker showing on which day and time what job will be executed

3.1 Baseline Scheduling

in ref. to Herroelen, Leus (2005) without any anticipation of variability

```
In [19]: import gurobipy as gp

# Create an environment with your Gurobi license
params = {
    'WLSACCESSID': '4a72e635-8c05-4d4a-b728-95a831b9ccdd',
    'WLSSECRET': 'f4ad3f05-2e2a-429b-a748-3e3180bc6c528',
    'LICENSEID': 'B06330'
}
env = gp.Env(params=params)
```

```
In [20]: from typing import List, Tuple
from gurobipy import Model, GRB

def generate_timetables(jobs: List[Tuple[int, int]], max_workers: int) -> None:
    """
    Generates timetables for a list of jobs, assigning workers to execute them. The goal is to minimize the
    number of workers needed, while ensuring that each job is executed exactly once, and that each worker ex-
    ecutes at most one job at any given time.
    """
```

```
    #param jobs: a list of tuples representing jobs to be executed. Each tuple should contain a job ID and t
    #executetime in 1/6 hours.
    #return: None
    """
```

```
    n_jobs = len(jobs)
    max_exec_time = max(jobs[i][1] for job in jobs)
    max_timesteps = 7 * 24 * 6 # One week in minutes
    max_workers = max_workers # Maximum number of workers allowed
    time_limit = 6000 # Time limit for the solver in seconds
    """
```

```
    # Create a Gurobi model
    baselineModel = Model(env)

    # Set a time limit for the solver
    baselineModel.setParam(GRB.Param.TimeLimit, time_limit)
```

```
    # Variables
    x = {} # Binary variables to indicate if worker i starts job j at time t
    y = {} # Binary variables to indicate if worker i is being used

    # Add variables to the model
    for i in range(max_workers):
        for j in range(n_jobs):
            y[i, j] = baselineModel.addVar(vtype=GRB.BINARY)
```

```
    # Constraints
    # Each job has to be executed once
    for j in range(n_jobs):
        job_length = min(jobs[j][1], max_timesteps)
        constraint_sum = sum(x[i, j, t] for i in range(max_workers) for t in range(max_timesteps) if t <= ma
        baselineModel.addConstr(constraint_sum == 1, f'job_{j}_executed_once')
```

```
    # Binary handling
    if max_timesteps - jobs[j][1] < 0:
        print(f"Warning: Execution time for job {j} exceeds max_timesteps.")

    # A worker can only execute one job simultaneously
    for i in range(max_workers):
        for t in range(max_timesteps):
            baselineModel.addConstr(sum(x[i, j, t] for j in range(n_jobs) for k in range(jobs[j][1]) if
```

```
    # Objective
    baselineModel.setObjective(sum(y[i] for i in range(max_workers)), GRB.MINIMIZE)

    # Solve the model
    baselineModel.optimize()
```

```
    # Print results and generate timetables
    if baselineModel.status == GRB.OPTIMAL:
        print("Optimal solution found.")
        print(f"Minimum number of workers: {int(baselineModel.objVal)}")
```

```
    # Generate timetables for each worker
    worker_timetables = {}

    for i in range(max_workers):
        worker_timetables[i] = []
        for j in range(n_jobs):
            for t in range(max_timesteps):
                if x[i, j, t] > 0.5:
                    worker_timetables[i].append((j, t, jobs[j][1]))
```

```
    # Sort by start time
    for worker, timetable in worker_timetables.items():
        timetable.sort(key=lambda x: x[1])
        print(f"Worker {worker} timetable:")
        for job, start, end in timetable:
            day_start = start // (24 * 6)
            hour_start = (start // 6) % 24
            minute_start = start % 6 * 10
```

```
            day_end = end // (24 * 6)
            hour_end = (end // 6) % 24
            minute_end = end % 6 * 10

            duration = jobs[job][1] + 10
            print(f"Job {job} (duration: {duration} minutes): Start at Day {day_start}, (hour_star
```

```
            else:
                print("No optimal solution found.")

    # Error handling
    if baselineModel.status == GRB.INFEASIBLE:
        print("Model is infeasible.")
        baselineModel.computeIIS()
        baselineModel.write("infeasible.ilp")
```

```
In [21]: # Test model
generate_timetables(jobs,3)
```


Set parameter TimeLimit to value 6000
Gurobi Optimizer version 10.0.1 build v10.0.1rc0 (win64)

CPU model: Intel(R) Xeon(R) Platinum 8171M CPU @ 2.60GHz, instruction set [SSE2] [AVX] [AVX2]
Thread count: 4 physical cores, 4 logical processors, using up to 4 threads

Academic license - for non-commercial use only - registered to nills.hellmann@stud-mail.uni-wuerzburg.de
Model fingerprint: 0x404940d5

Variable types: 0 continuous, 756003 integer (756003 binary)
Coefficient statistics:
Matrix range: [1e+00, 1e+00]
Objective range: [1e+00, 1e+00]
Bounds range: [1e+00, 1e+00]
RHS range: [1e+00, 1e+00]

Found heuristic solution: objective 6.0000000e+00
Presolve time: 3.18s
Presolved: 3274 rows, 756003 columns, 1515024 nonzeros
Variable types: 0 continuous, 756003 integer (756003 binary)

Root simplex log...

Iteration Objective Primal Inf. Dual Inf. Time
0 2.4801587e-01 0.000000e+00 0.000000e+00 5s

Use crossover to convert LP symmetric solution to basic solution...

Root crossover log...

754995 PPushes remaining with Pinf 0.0000000e+00 5s
0 PPushes remaining with Pinf 0.0000000e+00
0 DPushes remaining with Dinf 0.0000000e+00 9s

Push phase complete: Pinf 0.0000000e+00, Dinf 6.7902775e-15 9s

Root simplex log...

Iteration Objective Primal Inf. Dual Inf. Time
754998 2.4801587e-01 0.000000e+00 0.000000e+00 9s

Root relaxation: objective 2.480159e-01, 754998 iterations, 5.33 seconds (3.59 work units)

Nodes	Current Node	Objective Bounds	BestBd	Gap	IT/Node/Time
Expl Unexpl	Obj Depth IntInf	Incumbent			
0	0	0.24802	0.24802	91.7%	- 47s
H	0	0	1.0000000	0.24802	75.2% - 48s
0	0	0.24802	0.3272	1.000000	0.24802 75.2% - 159s

Explored 1 nodes (76417 simplex iterations) in 48.36 seconds (34.77 work units)
Thread count was 4 (of 4 available processors)

Solution count: 2: 1: 3

Optimal solution found (tolerance 1.00e-04)
Best objective 1.000000000000e+00, best bound 1.000000000000e+00, gap 0.0000%

Optimal solution found.
Minimum number of workers: 1
Worker: 1

Job 235 (duration: 10 minutes): Start at Day 0, 00:00 - End at Day 0, 00:10
Job 136 (duration: 10 minutes): Start at Day 0, 00:10 - End at Day 0, 00:20
Job 228 (duration: 10 minutes): Start at Day 0, 00:20 - End at Day 0, 02:40
Job 108 (duration: 10 minutes): Start at Day 0, 00:40 - End at Day 0, 00:50
Job 74 (duration: 10 minutes): Start at Day 0, 00:50 - End at Day 0, 01:00
Job 206 (duration: 10 minutes): Start at Day 0, 01:00 - End at Day 0, 01:10
Job 218 (duration: 10 minutes): Start at Day 0, 01:10 - End at Day 0, 01:20
Job 190 (duration: 10 minutes): Start at Day 0, 01:20 - End at Day 0, 01:30
Job 209 (duration: 10 minutes): Start at Day 0, 01:30 - End at Day 0, 01:40
Job 147 (duration: 10 minutes): Start at Day 0, 01:40 - End at Day 0, 01:50
Job 89 (duration: 10 minutes): Start at Day 0, 02:00 - End at Day 0, 02:10
Job 195 (duration: 10 minutes): Start at Day 0, 02:10 - End at Day 0, 02:20
Job 157 (duration: 10 minutes): Start at Day 0, 02:20 - End at Day 0, 02:30
Job 101 (duration: 10 minutes): Start at Day 0, 02:30 - End at Day 0, 02:40
Job 223 (duration: 10 minutes): Start at Day 0, 02:40 - End at Day 0, 02:50
Job 29 (duration: 10 minutes): Start at Day 0, 02:50 - End at Day 0, 03:00
Job 201 (duration: 10 minutes): Start at Day 0, 03:00 - End at Day 0, 03:10
Job 224 (duration: 10 minutes): Start at Day 0, 03:10 - End at Day 0, 03:20
Job 65 (duration: 10 minutes): Start at Day 0, 03:20 - End at Day 0, 03:30
Job 126 (duration: 10 minutes): Start at Day 0, 03:30 - End at Day 0, 03:40
Job 181 (duration: 10 minutes): Start at Day 0, 03:40 - End at Day 0, 03:50
Job 115 (duration: 10 minutes): Start at Day 0, 03:50 - End at Day 0, 04:00
Job 179 (duration: 10 minutes): Start at Day 0, 04:00 - End at Day 0, 04:10
Job 161 (duration: 10 minutes): Start at Day 0, 04:10 - End at Day 0, 04:20
Job 248 (duration: 10 minutes): Start at Day 0, 04:20 - End at Day 0, 04:30
Job 163 (duration: 10 minutes): Start at Day 0, 04:30 - End at Day 0, 04:40
Job 77 (duration: 10 minutes): Start at Day 0, 04:40 - End at Day 0, 04:50
Job 202 (duration: 10 minutes): Start at Day 0, 04:50 - End at Day 0, 05:00
Job 229 (duration: 10 minutes): Start at Day 0, 05:00 - End at Day 0, 05:10
Job 15 (duration: 10 minutes): Start at Day 0, 05:10 - End at Day 0, 05:20
Job 228 (duration: 10 minutes): Start at Day 0, 05:20 - End at Day 0, 05:30
Job 238 (duration: 10 minutes): Start at Day 0, 05:30 - End at Day 0, 05:40
Job 171 (duration: 10 minutes): Start at Day 0, 05:40 - End at Day 0, 05:50
Job 35 (duration: 10 minutes): Start at Day 0, 05:50 - End at Day 0, 06:00
Job 234 (duration: 10 minutes): Start at Day 0, 06:00 - End at Day 0, 06:10
Job 139 (duration: 10 minutes): Start at Day 0, 06:10 - End at Day 0, 06:30
Job 53 (duration: 10 minutes): Start at Day 0, 06:30 - End at Day 0, 06:40
Job 189 (duration: 10 minutes): Start at Day 0, 06:40 - End at Day 0, 06:50
Job 69 (duration: 10 minutes): Start at Day 0, 06:50 - End at Day 0, 07:00
Job 196 (duration: 10 minutes): Start at Day 0, 07:00 - End at Day 0, 07:10
Job 26 (duration: 10 minutes): Start at Day 0, 07:10 - End at Day 0, 07:20
Job 216 (duration: 10 minutes): Start at Day 0, 07:20 - End at Day 0, 07:30
Job 103 (duration: 10 minutes): Start at Day 0, 07:30 - End at Day 0, 07:40
Job 247 (duration: 10 minutes): Start at Day 0, 07:40 - End at Day 0, 07:50
Job 174 (duration: 10 minutes): Start at Day 0, 08:00 - End at Day 0, 08:10
Job 174 (duration: 10 minutes): Start at Day 0, 08:10 - End at Day 0, 08:20
Job 182 (duration: 10 minutes): Start at Day 0, 08:20 - End at Day 0, 08:30
Job 230 (duration: 10 minutes): Start at Day 0, 08:30 - End at Day 0, 08:40
Job 122 (duration: 10 minutes): Start at Day 0, 08:50 - End at Day 0, 09:00
Job 191 (duration: 10 minutes): Start at Day 0, 09:00 - End at Day 0, 09:10
Job 176 (duration: 10 minutes): Start at Day 0, 09:10 - End at Day 0, 09:20
Job 70 (duration: 10 minutes): Start at Day 0, 09:20 - End at Day 0, 09:30
Job 134 (duration: 10 minutes): Start at Day 0, 09:30 - End at Day 0, 09:40
Job 245 (duration: 10 minutes): Start at Day 0, 09:40 - End at Day 0, 09:50
Job 107 (duration: 10 minutes): Start at Day 0, 09:50 - End at Day 0, 10:00
Job 68 (duration: 10 minutes): Start at Day 0, 10:00 - End at Day 0, 10:10
Job 185 (duration: 10 minutes): Start at Day 0, 10:20 - End at Day 0, 10:30
Job 154 (duration: 10 minutes): Start at Day 0, 10:30 - End at Day 0, 10:40
Job 90 (duration: 10 minutes): Start at Day 0, 10:40 - End at Day 0, 10:50
Job 216 (duration: 10 minutes): Start at Day 0, 11:00 - End at Day 0, 11:10
Job 111 (duration: 10 minutes): Start at Day 0, 11:10 - End at Day 0, 11:20
Job 184 (duration: 10 minutes): Start at Day 0, 11:20 - End at Day 0, 11:30
Job 93 (duration: 10 minutes): Start at Day 0, 11:30 - End at Day 0, 11:50
Job 142 (duration: 10 minutes): Start at Day 0, 11:50 - End at Day 0, 12:00
Job 244 (duration: 10 minutes): Start at Day 0, 12:00 - End at Day 0, 12:10
Job 61 (duration: 10 minutes): Start at Day 0, 12:10 - End at Day 0, 12:20
Job 177 (duration: 10 minutes): Start at Day 0, 12:20 - End at Day 0, 12:30
Job 241 (duration: 10 minutes): Start at Day 0, 12:30 - End at Day 0, 12:40
Job 187 (duration: 10 minutes): Start at Day 0, 12:40 - End at Day 0, 12:50
Job 23 (duration: 10 minutes): Start at Day 0, 13:00 - End at Day 0, 13:10
Job 213 (duration: 10 minutes): Start at Day 0, 13:10 - End at Day 0, 13:30
Job 20 (duration: 10 minutes): Start at Day 0, 13:30 - End at Day 0, 13:40
Job 228 (duration: 10 minutes): Start at Day 0, 13:40 - End at Day 0, 13:50
Job 237 (duration: 10 minutes): Start at Day 0, 13:50 - End at Day 0, 14:00
Job 119 (duration: 10 minutes): Start at Day 0, 14:00 - End at Day 0, 14:10
Job 51 (duration: 10 minutes): Start at Day 0, 14:10 - End at Day 0, 14:20
Job 45 (duration: 10 minutes): Start at Day 0, 14:20 - End at Day 0, 14:30
Job 141 (duration: 10 minutes): Start at Day 0, 14:30 - End at Day 0, 14:40
Job 190 (duration: 10 minutes): Start at Day 0, 14:40 - End at Day 0, 14:50
Job 118 (duration: 10 minutes): Start at Day 0, 14:50 - End at Day 0, 15:00
Job 192 (duration: 10 minutes): Start at Day 0, 15:00 - End at Day 0, 15:10
Job 79 (duration: 10 minutes): Start at Day 0, 15:10 - End at Day 0, 15:20
Job 109 (duration: 10 minutes): Start at Day 0, 15:20 - End at Day 0, 15:30
Job 119 (duration: 10 minutes): Start at Day 0, 15:40 - End at Day 0, 15:50
Job 41 (duration: 10 minutes): Start at Day 0, 15:50 - End at Day 0, 16:00
Job 219 (duration: 10 minutes): Start at Day 0, 16:00 - End at Day 0, 16:10
Job 67 (duration: 10 minutes): Start at Day 0, 16:10 - End at Day 0, 16:20
Job 151 (duration: 10 minutes): Start at Day 0, 16:20 - End at Day 0, 16:40
Job 159 (duration: 10 minutes): Start at Day 0, 16:40 - End at Day 0, 16:50
Job 80 (duration: 10 minutes): Start at Day 0, 16:50 - End at Day 0, 17:00
Job 158 (duration: 10 minutes): Start at Day 0, 17:00 - End at Day 0, 17:10
Job 159 (duration: 10 minutes): Start at Day 0, 17:10 - End at Day 0, 17:30
Job 215 (duration: 10 minutes): Start at Day 0, 17:30 - End at Day 0, 17:40
Job 156 (duration: 10 minutes): Start at Day 0, 17:50 - End at Day 0, 18:00
Job 102 (duration: 10 minutes): Start at Day 0, 18:00 - End at Day 0, 18:10
Job 129 (duration: 10 minutes): Start at Day 0, 18:10 - End at Day 0, 18:30
Job 128 (duration: 10 minutes): Start at Day 0, 18:30 - End at Day 0, 18:40
Job 212 (duration: 10 minutes): Start at Day 0, 18:40 - End at Day 0, 19:30
Job 22 (duration: 10 minutes): Start at Day 0, 19:30 - End at Day 0, 19:50
Job 39 (duration: 10 minutes): Start at Day 0, 19:50 - End at Day 0, 20:10
Job 22 (duration: 10 minutes): Start at Day 0, 20:10 - End at Day 0, 21:20
Job 99 (duration: 10 minutes): Start at Day 0, 21:20 - End at Day 0, 21:30
Job 125 (duration: 10 minutes): Start at Day 0, 22:30 - End at Day 0, 22:40
Job 238 (duration: 10 minutes): Start at Day 0, 22:40 - End at Day 0, 22:50
Job 135 (duration: 10 minutes): Start at Day 0, 22:50 - End at Day 0, 23:00
Job 94 (duration: 10 minutes): Start at Day 0, 23:00 - End at Day 0, 23:00
Job 114 (duration: 10 minutes): Start at Day 0, 23:10 - End at Day 0, 23:20
Job 3 (duration: 10 minutes): Start at Day 0, 23:20 - End at Day 0, 23:30
Job 243 (duration: 10 minutes): Start at Day 0, 23:30 - End at Day 0, 23:40
Job 63 (duration: 10 minutes): Start at Day 0, 23:50 - End at Day 1, 00:00
Job 91 (duration: 10 minutes): Start at Day 1, 00:00 - End at Day 1, 00:10
Job 24 (duration: 10 minutes): Start at Day 1, 00:10 - End at Day 1, 00:20
Job 198 (duration: 10 minutes): Start at Day 1, 00:20 - End at Day 1, 00:50
Job 194 (duration: 10 minutes): Start at Day 1, 00:20 - End at Day 1, 02:30
Job 240 (duration: 10 minutes): Start at Day 1, 02:30 - End at Day 1, 02:40
Job 50 (duration: 10 minutes): Start at Day 1, 02:40 - End at Day 1, 02:50
Job 99 (duration: 10 minutes): Start at Day 1, 02:50 - End at Day 1, 02:30
Job 205 (duration: 10 minutes): Start at Day 1, 02:40 - End at Day 1, 02:50
Job 46 (duration: 10 minutes): Start at Day 1, 04:00 - End at Day 1, 04:10
Job 204 (duration: 10 minutes): Start at Day 1, 04:10 - End at Day 1, 04:50
Job 132 (duration: 10 minutes): Start at Day 1, 04:00 - End at Day 1, 05:10
Job 208 (duration: 10 minutes): Start at Day 1, 05:10 - End at Day 1, 05:20
Job 173 (duration: 10 minutes): Start at Day 1, 05:50 - End at Day 1, 06:00
Job 217 (duration: 10 minutes): Start at Day 1, 06:00 - End at Day 1, 06:10
Job 127 (duration: 10 minutes): Start at Day 1, 06:20 - End at Day 1, 06:20
Job 175 (duration: 10 minutes): Start at Day 1, 06:20 - End at Day 1, 07:20
Job 203 (duration: 10 minutes): Start at Day 1, 07:20 - End at Day 1, 07:50
Job 196 (duration: 10 minutes): Start at Day 1, 07:50 - End at Day 1, 08:00
Job 224 (duration: 10 minutes): Start at Day 1, 08:30 - End at Day 1, 09:40
Job 226 (duration: 10 minutes): Start at Day 1, 09:40 - End at Day 1, 09:50
Job 98 (duration: 10 minutes): Start at Day 1, 09:50 - End at Day 1, 09:50
Job 222 (duration: 10 minutes): Start at Day 1, 10:00 - End at Day 1, 10:10
Job 148 (duration: 10 minutes): Start at Day 1, 10:20 - End at Day 1, 10:30
Job 106 (duration: 10 minutes): Start at Day 1, 10:30 - End at Day 1, 10:40
Job 180 (duration: 10 minutes): Start at Day 1, 10:40 - End at Day 1, 10:50
Job 52 (duration: 10 minutes): Start at Day 1, 10:50 - End at Day 1, 11:00
Job 137 (duration: 10 minutes): Start at Day 1, 11:00 - End at Day 1, 11:10
Job 157 (duration: 10 minutes): Start at Day 1, 11:10 - End at Day 1, 11:30
Job 130 (duration: 10 minutes): Start at Day 1, 11:30 - End at Day 1, 11:40
Job 166 (duration: 10 minutes): Start at Day 1, 11:40 - End at Day 1, 11:50
Job 57 (duration: 10 minutes): Start at Day 1, 12:10 - End at Day 1, 12:10
Job 62 (duration: 10 minutes): Start at Day 1, 12:10 - End at Day 1, 12:10
Job 123 (duration: 10 minutes): Start at Day 1, 14:10 - End at Day 1, 14:20
Job 118 (duration: 10 minutes): Start at Day 1, 14:20 - End at Day 1, 14:30
Job 155 (duration: 10 minutes): Start at Day 1, 16:00 - End at Day 1, 16:10
Job 16 (duration: 10 minutes): Start at Day 1, 16:10 - End at Day 1, 16:20
Job 217 (duration: 10 minutes): Start at Day 1, 16:20 - End at Day 1, 16:30
Job 92 (duration: 10 minutes): Start at Day 1, 17:20 - End at Day 1, 17:20
Job 177 (duration: 10 minutes): Start at Day 1, 17:20 - End at Day 1, 17:30
Job 124 (duration: 10 minutes): Start at Day 1, 17:30 - End at Day 1, 17:40
Job 217 (duration: 10 minutes): Start at Day 1, 17:30 - End at Day 1, 17:40
Job 239 (duration: 10 minutes): Start at Day 1, 19:00 - End at Day 1, 19:10
Job 21 (duration: 10 minutes): Start at Day 1, 19:00 - End at Day 1, 19:40
Job 124 (duration: 10 minutes): Start at Day 1, 19:50 - End at Day 1, 20:00
Job 164 (duration: 10 minutes): Start at Day 1, 19:50 - End at Day 1, 20:00
Job 19 (duration: 10 minutes): Start at Day 1, 20:10 - End at Day 1, 20:20
Job 159 (duration: 10 minutes): Start at Day 1, 20:20 - End at Day 1, 21:30
Job 168 (duration: 10 minutes): Start at Day 1, 20:40 - End at Day 1, 20:50
Job 227 (duration: 10 minutes): Start at Day 1, 21:10 - End at Day 1, 21:20
Job 81 (duration: 10 minutes): Start at Day 1, 21:30 - End at Day 1, 21:40
Job 30 (duration: 10 minutes): Start at Day 1, 22:00 - End at Day 1, 22:10
Job 213 (duration: 10 minutes): Start at Day 1, 22:10 - End at Day 1, 22:30
Job 84 (duration: 10 minutes): Start at Day 1, 22:10 - End at Day 1, 23:20
Job 213 (duration: 10 minutes): Start at Day 2, 00:10 - End at Day 2, 00:20
Job 98 (duration: 10 minutes): Start at Day 2, 00:30 - End at Day 2, 00:40
Job 78 (duration: 10 minutes): Start at Day 2, 01:10 - End at Day 2, 01:30
Job 140 (duration: 10 minutes): Start at Day 2, 01:20 - End at Day 2, 01:30
Job 76 (duration: 10 minutes): Start at Day 2, 01:40 - End at Day 2, 02:50
Job 180 (duration: 10 minutes): Start at Day 2, 02:10 - End at Day 2, 02:50
Job 214 (duration: 10 minutes): Start at Day 2, 03:10 - End at Day 2, 03:20
Job 26 (duration: 10 minutes): Start at Day 2, 03:20 - End at Day 2, 03:30
Job 175 (duration: 10 minutes): Start at Day 2, 03:30 - End at Day 2, 04:00
Job 121 (duration: 10 minutes): Start at Day 2, 04:00 - End at Day 2, 04:10
Job 10 (duration: 10 minutes): Start at Day 2, 04:50 - End at Day 2, 05:10
Job 36 (duration: 10 minutes): Start at Day 2, 05:10 - End at Day 2, 05:10
Job 203 (duration: 10 minutes): Start at Day 2, 06:00 - End at Day 2, 06:10
Job 48 (duration: 10 minutes): Start at Day 2, 07:20 - End at Day 2, 07:20
Job 98 (duration: 10 minutes): Start at Day 2, 08:10 - End at Day 2, 08:20
Job 176 (duration: 10 minutes): Start at Day 2, 08:20 - End at Day 2, 08:30
Job 85 (duration: 10 minutes): Start at Day 2, 09:00 - End at Day 2, 09:10
Job 133 (duration: 10 minutes): Start at Day 2, 09:30 - End at Day 2, 09:40
Job 27 (duration: 10 minutes): Start at Day 2, 09:50 - End at Day 2, 09:50
Job 169 (duration: 10 minutes): Start at Day 2, 09:50 - End at Day 2, 11:10
Job 54 (duration: 10 minutes): Start at Day 2, 11:50 - End at Day 2, 11:30
Job 118 (duration: 10 minutes): Start at Day 2, 13:50 - End at Day 2, 14:00
Job 144 (duration: 10 minutes): Start at Day 2, 13:50 - End at Day 2, 14:00
Job 37 (duration: 10 minutes): Start at Day 2, 14:00 - End at Day 2, 14:10
Job 71 (duration: 10 minutes): Start at Day 2, 14:10 - End at Day 2, 14:20
Job 47 (duration: 10 minutes): Start at Day 2, 14:30 - End at Day 2, 14:40
Job 249 (duration: 10 minutes): Start at Day 2, 15:00 - End at Day 2, 15:10
Job 245 (duration: 10 minutes): Start at Day 2, 16:00 - End at Day 2, 16:10
Job 55 (duration: 10 minutes): Start at Day 2, 16:00 - End at Day 2, 16:10
Job 143 (duration: 10 minutes): Start at Day 2, 17:10 - End at Day 2, 17:20
Job 75 (duration: 10 minutes): Start at Day 2, 17:20 - End at Day 2, 17:30
Job 96 (duration: 10 minutes): Start at Day 2, 17:30 - End at Day 2, 18:00
Job 49 (duration: 10 minutes): Start at Day 2, 18:00 - End at Day 2, 18:40
Job 153 (duration: 10 minutes): Start at Day 2, 18:50 - End at Day 2, 19:00
Job 113 (duration: 10 minutes): Start at Day 2, 19:10 - End at Day 2, 19:20
Job 33 (duration: 10 minutes): Start at Day 2, 19:20 - End at Day 2, 19:30
Job 181 (duration: 10 minutes): Start at Day 2, 21:20 - End at Day 2, 21:30
Job 34 (duration: 10 minutes): Start at Day 2, 21:30 - End at Day 2, 22:10
Job 117 (duration: 10 minutes): Start at Day 2, 22:00 - End at Day 2, 22:10
Job 38 (duration: 10 minutes): Start at Day 2, 22:50 - End at Day 3, 00:00
Job 188 (duration: 10 minutes): Start at Day 3, 00:20 - End at Day 3, 00:30
Job 42 (duration: 10 minutes): Start at Day 3, 00:40 - End at Day 3, 00:50
Job 149 (duration: 10 minutes): Start at Day 3, 01:30 - End at Day 3, 01:40
Job 11 (duration: 10 minutes): Start at Day 3, 01:40 - End at Day 3, 01:50
Job 43 (duration: 10 minutes): Start at Day 3, 01:50 - End at Day 3, 02:00
Job 7 (duration: 10 minutes): Start at Day 3, 02:50 - End at Day 3, 03:00
Job 83 (duration: 10 minutes): Start at Day 3, 04:20 - End at Day 3, 04:30
Job 56 (duration: 10 minutes): Start at Day 3, 04:40 - End at Day 3, 04:50
Job 278 (duration: 10 minutes): Start at Day 3, 06:30 - End at Day 3, 06:40
Job 146 (duration: 10 minutes): Start at Day 3, 07:10 - End at Day 3, 07:20
Job 71 (duration: 10 minutes): Start at Day 3, 08:30 - End at Day 3, 08:40
Job 64 (duration: 10 minutes): Start at Day 3, 09:30 - End at Day 3, 09:40
Job 59 (duration: 10 minutes): Start at Day 3, 11:50 - End at Day 3, 12:00
Job 225 (duration: 10 minutes): Start at Day 3, 12:50 - End at Day 3, 13:00
Job 129 (duration: 10 minutes): Start at Day 3, 13:50 - End at Day 3, 14:00
Job 128 (duration: 10 minutes): Start at Day 3, 19:00 - End at Day 3, 19:10
Job 14 (duration: 10 minutes): Start at Day 3, 19:00 - End at Day 3, 20:30
Job 5 (duration: 10 minutes): Start at Day 3, 20:30 - End at Day 3, 20:30
Job 160 (duration: 10 minutes): Start at Day 3, 22:20 - End at Day 3, 22:30
Job 97 (duration: 10 minutes): Start at Day 3, 22:30 - End at Day 3, 23:40
Job 112 (duration: 10 minutes): Start at Day 3, 23:40 - End at Day 3, 23:50
Job 60 (duration: 10 minutes): Start at Day 4, 01:00 - End at Day 4, 01:10
Job 31 (duration: 10 minutes): Start at Day 4, 02:00 - End at Day 4, 02:10
Job 58 (duration: 10 minutes): Start at Day 4, 03:40 - End at Day 4, 03:50
Job 13 (duration: 10 minutes): Start at Day 4, 03:50 - End at Day 4, 04:00
Job 9 (duration: 10 minutes): Start at Day 4, 07:20 - End at Day 4, 07:30
Job 73 (duration: 10 minutes): Start at Day 4, 15:50 - End at Day 4, 16:00
Job 22 (duration: 10 minutes): Start at Day 4, 16:00 - End at Day 4, 16:10
Job 24 (duration: 10 minutes): Start at Day 4, 03:30 - End at Day 4, 03:40
Job 112 (duration: 10 minutes): Start at Day 5, 05:40 - End at Day 5, 05:50
Job 0 (duration: 10 minutes): Start at Day 5, 07:10 - End at Day 5, 07:20
Job 6 (duration: 10 minutes): Start at Day 5, 13:50 - End at Day 5, 14:00
Job 2 (duration: 10 minutes): Start at Day 5, 14:00 - End at Day 5, 14:10
Job 3 (duration: 10 minutes): Start at Day 5, 20:40 - End at Day 5, 20:50
Job 1 (duration: 10 minutes): Start at Day 6, 10:20 - End at Day 6, 10:30

3.2 Smart Stochastic scheduling (SSS)

3.2.1 Loss function

We use the Mean Squared Error (MSE) because it's sensitive to large errors and works well for continuous values like runtimes.

loss_mse = the decision error induced by a prediction

```
In [22]: import numpy as np
import scipy.stats as st

def loss_function(data, confidence_level):
    """
    This function can be used to determine confident runtimes for scheduling processes in a calendar,
    taking into account the uncertainty in the predictions based on the model's MSE and the desired confidence.
    The resulting confident runtimes represent the upper bounds of the confidence intervals for the prediction
    ensuring that the actual runtimes are likely to be within the specified confidence level below these con-
    fidence levels.

    :param jobs: A dataframe with the future jobs.
    :param confidence_level: Input confidence level.
    :return: A list of the confident runtimes for each job.
    """

    # Assume your predictions and the MSE are stored in the following variables:
    predictions = data["ExecTime"].astype(int)
    mse = model_mse

    # Calculate the standard deviation of the prediction errors (RMSE)
    rmse = np.sqrt(mse)

    # Assign the confidence level
    z_value = st.norm.ppf((1 + confidence_level) / 2)

    # Calculate the confidence interval for each prediction
    confident_runtimes = predictions + z_value * rmse

    return confident_runtimes

# Call function
confident_runtimes = loss_function(future_jobs, 0.99)
print(confident_runtimes)
```

0 108.500641
1 113.500641
2 57.500641
3 49.500641
4 87.500641
...
245 49.500641
246 49.500641
247 49.500641
248 49.500641
249 49.500641
Name: ExecTime, Length: 250, dtype: float64

Create a graph to visualize the relation between confidence levels and runtimes.

```
In [23]: # Define a range of confidence levels from 58% to 99.999%
confidence_levels = np.linspace(0.59, 100)

# Calculate confident runtimes for each confidence level
runtimes = [loss_function(future_jobs, cl) for cl in confidence_levels]

# Calculate the average confident runtime for each confidence level
average_runtimes = [np.mean(rt) for rt in runtimes]
```

```
# Set the size of the plot (width, height) in inches
plt.figure(figsize=(9, 3))

# Plot the sensitivity analysis
plt.plot(confidence_levels + 100, average_runtimes)
```

```
# Customize x-axis tick labels
x_ticks = np.arange(0, 101, 10) # Define tick positions
x_tick_labels = [f'{t} %' for t in x_ticks] # Add a symbol to tick labels
plt.xticks(x_ticks, x_tick_labels)
```

```
# Customize y-axis tick labels
y_ticks = np.arange(np.ceil(np.min(average_runtimes)), np.ceil(np.max(average_runtimes)) + 5, 6) # Define t
y_tick_labels = [f'{t} s' for t in y_ticks] # Add a 's' unit to tick labels
plt.yticks(y_ticks, y_tick_labels)
```

```
plt.xlabel("Confidence Level")
plt.ylabel("Average Confident Runtime")
plt.title("Sensitivity Analysis of the Loss Function")
```

```
# Save the plot as
```


[illegible]

Optimal solution found.
Minimum number of workers for: 97.0648275862683% completion: 22860
Ausführungszeit: 369.1973757832801 min
Set parameter TimeLimit to value 6000
Gurobi Optimizer version 10.0.1 build v10.0.1.rc0 (win64)

CPU model: Intel(R) Xeon(R) Platinum 8171M CPU @ 2.60GHz, instruction set [SSE2]AVX[AVX2]AVXS12
Thread count: 4 physical cores, 4 logical processors, using up to 4 threads

Academic license - for non-commercial use only - registered to nlls.hellmann@stud-mail.uni-wuerzburg.de
Optimize a model with 6298 rows, 1512006 columns and 3038048 nonzeros

Model fingerprint: 0x4d47dda5

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Coefficient statistics:

Matrix range [1e+00, 1e+00]

Objective range [1e+00, 1e+00]

Bounds range [1e+00, 1e+00]

RHS range [1e+00, 1e+00]

Found heuristic solution: objective 6.00000000

Presolve removed 0 rows and 0 columns (presolve time = 5s) ...

Presolve time: 6.51s

Presolved: 6298 rows, 1512006 columns, 3038048 nonzeros

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
0	2.4801587e-01	0.000000e+00	0.000000e+00	10s

Use crossover to convert LP symmetric solution to basic solution...

Root crossover log...

1510998	PPushes remaining with Pinf 0.0000000e+00	10s
1485750	PPushes remaining with Pinf 0.0000000e+00	10s
880904	PPushes remaining with Pinf 0.0000000e+00	15s
241180	PPushes remaining with Pinf 0.0000000e+00	20s
0	PPushes remaining with Pinf 0.0000000e+00	22s
0	DPushes remaining with Dinf 0.0000000e+00	22s

Push phase complete: Pinf 0.0000000e+00, Dinf 4.3723705e-15

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
1511001	2.4801587e-01	0.000000e+00	0.000000e+00	22s

Root relaxation: objective 2.480159e-01, 1511001 iterations, 14.55 seconds (11.20 work units)

Nodes	Current Node	Objective Bounds	Work					
Expl Unexpl	Obj Depth IntInf	Incumbent BestBd Gap	It/Node Time					
0	0	0.24802	0.6296	6.000000	0.24802	95.9%	-	156s
H	0	0	0	1.0000000	0.24802	75.2%	-	157s
0	0	0.24802	0.6296	1.00000	0.24802	75.2%	-	157s

Explored 1 nodes (1513446 simplex iterations) in 157.57 seconds (83.79 work units)

Thread count was 4 (of 4 available processors)

Solution count 2: 1 6

Optimal solution found (tolerance 1.00e-04)

Best objective 1.000000000000e+00, best bound 1.000000000000e+00, gap 0.0000%

Optimal solution found.

Minimum number of workers for: 97.2146951724137% completion: 22151

Ausführungszeit: 370.72909358762556 min

Set parameter TimeLimit to value 6000

Gurobi Optimizer version 10.0.1 build v10.0.1.rc0 (win64)

CPU model: Intel(R) Xeon(R) Platinum 8171M CPU @ 2.60GHz, instruction set [SSE2]AVX[AVX2]AVXS12
Thread count: 4 physical cores, 4 logical processors, using up to 4 threads

Academic license - for non-commercial use only - registered to nlls.hellmann@stud-mail.uni-wuerzburg.de
Optimize a model with 6298 rows, 1512006 columns and 3038048 nonzeros

Model fingerprint: 0x4d47dda5

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Coefficient statistics:

Matrix range [1e+00, 1e+00]

Objective range [1e+00, 1e+00]

Bounds range [1e+00, 1e+00]

RHS range [1e+00, 1e+00]

Found heuristic solution: objective 6.00000000

Presolve removed 0 rows and 0 columns (presolve time = 5s) ...

Presolve time: 6.62s

Presolved: 6298 rows, 1512006 columns, 3038048 nonzeros

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
0	2.4801587e-01	0.000000e+00	0.000000e+00	10s

Use crossover to convert LP symmetric solution to basic solution...

Root crossover log...

1510998	PPushes remaining with Pinf 0.0000000e+00	10s
1504800	PPushes remaining with Pinf 0.0000000e+00	10s
818914	PPushes remaining with Pinf 0.0000000e+00	15s
245361	PPushes remaining with Pinf 0.0000000e+00	20s
0	PPushes remaining with Pinf 0.0000000e+00	22s
0	DPushes remaining with Dinf 0.0000000e+00	22s

Push phase complete: Pinf 0.0000000e+00, Dinf 4.3723705e-15

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
1511001	2.4801587e-01	0.000000e+00	0.000000e+00	22s

Root relaxation: objective 2.480159e-01, 1511001 iterations, 14.55 seconds (11.20 work units)

Nodes	Current Node	Objective Bounds	Work					
Expl Unexpl	Obj Depth IntInf	Incumbent BestBd Gap	It/Node Time					
0	0	0.24802	0.6296	6.000000	0.24802	95.9%	-	157s
H	0	0	0	1.0000000	0.24802	75.2%	-	158s
0	0	0.24802	0.6296	1.00000	0.24802	75.2%	-	158s

Explored 1 nodes (1513446 simplex iterations) in 158.90 seconds (83.79 work units)

Thread count was 4 (of 4 available processors)

Solution count 2: 1 6

Optimal solution found (tolerance 1.00e-04)

Best objective 1.000000000000e+00, best bound 1.000000000000e+00, gap 0.0000%

Optimal solution found.

Minimum number of workers for: 97.5810344827586% completion: 22247

Ausführungszeit: 372.49538020072136 min

Set parameter TimeLimit to value 6000

Gurobi Optimizer version 10.0.1 build v10.0.1.rc0 (win64)

CPU model: Intel(R) Xeon(R) Platinum 8171M CPU @ 2.60GHz, instruction set [SSE2]AVX[AVX2]AVXS12
Thread count: 4 physical cores, 4 logical processors, using up to 4 threads

Academic license - for non-commercial use only - registered to nlls.hellmann@stud-mail.uni-wuerzburg.de
Optimize a model with 6298 rows, 1512006 columns and 3038048 nonzeros

Model fingerprint: 0x4d47dda5

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Coefficient statistics:

Matrix range [1e+00, 1e+00]

Objective range [1e+00, 1e+00]

Bounds range [1e+00, 1e+00]

RHS range [1e+00, 1e+00]

Found heuristic solution: objective 6.00000000

Presolve removed 0 rows and 0 columns (presolve time = 5s) ...

Presolve time: 6.55s

Presolved: 6298 rows, 1512006 columns, 3038048 nonzeros

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
0	2.4801587e-01	0.000000e+00	0.000000e+00	10s

Use crossover to convert LP symmetric solution to basic solution...

Root crossover log...

1510998	PPushes remaining with Pinf 0.0000000e+00	10s
1475939	PPushes remaining with Pinf 0.0000000e+00	10s
800904	PPushes remaining with Pinf 0.0000000e+00	15s
233605	PPushes remaining with Pinf 0.0000000e+00	20s
0	PPushes remaining with Pinf 0.0000000e+00	22s
0	DPushes remaining with Dinf 0.0000000e+00	22s

Push phase complete: Pinf 0.0000000e+00, Dinf 4.3723705e-15

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
1511001	2.4801587e-01	0.000000e+00	0.000000e+00	22s

Root relaxation: objective 2.480159e-01, 1511001 iterations, 14.42 seconds (11.20 work units)

Nodes	Current Node	Objective Bounds	Work					
Expl Unexpl	Obj Depth IntInf	Incumbent BestBd Gap	It/Node Time					
0	0	0.24802	0.6296	6.000000	0.24802	95.9%	-	156s
H	0	0	0	1.0000000	0.24802	75.2%	-	157s
0	0	0.24802	0.6296	1.00000	0.24802	75.2%	-	157s

Explored 1 nodes (1513446 simplex iterations) in 157.64 seconds (83.79 work units)

Thread count was 4 (of 4 available processors)

Solution count 2: 1 6

Optimal solution found (tolerance 1.00e-04)

Best objective 1.000000000000e+00, best bound 1.000000000000e+00, gap 0.0000%

Optimal solution found.

Minimum number of workers for: 97.5810344827586% completion: 22249

Ausführungszeit: 372.38012210721276 min

Set parameter TimeLimit to value 6000

Gurobi Optimizer version 10.0.1 build v10.0.1.rc0 (win64)

CPU model: Intel(R) Xeon(R) Platinum 8171M CPU @ 2.60GHz, instruction set [SSE2]AVX[AVX2]AVXS12
Thread count: 4 physical cores, 4 logical processors, using up to 4 threads

Academic license - for non-commercial use only - registered to nlls.hellmann@stud-mail.uni-wuerzburg.de
Optimize a model with 6298 rows, 1512006 columns and 3038048 nonzeros

Model fingerprint: 0x4d47dda5

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Coefficient statistics:

Matrix range [1e+00, 1e+00]

Objective range [1e+00, 1e+00]

Bounds range [1e+00, 1e+00]

RHS range [1e+00, 1e+00]

Found heuristic solution: objective 6.00000000

Presolve removed 0 rows and 0 columns (presolve time = 5s) ...

Presolve time: 6.47s

Presolved: 6298 rows, 1512006 columns, 3038048 nonzeros

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
0	2.4801587e-01	0.000000e+00	0.000000e+00	10s

Use crossover to convert LP symmetric solution to basic solution...

Root crossover log...

1510998	PPushes remaining with Pinf 0.0000000e+00	10s
1475939	PPushes remaining with Pinf 0.0000000e+00	10s
800904	PPushes remaining with Pinf 0.0000000e+00	15s
233790	PPushes remaining with Pinf 0.0000000e+00	20s
0	PPushes remaining with Pinf 0.0000000e+00	22s
0	DPushes remaining with Dinf 0.0000000e+00	22s

Push phase complete: Pinf 0.0000000e+00, Dinf 4.3723705e-15

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
1511001	2.4801587e-01	0.000000e+00	0.000000e+00	22s

Root relaxation: objective 2.480159e-01, 1511001 iterations, 14.57 seconds (11.20 work units)

Nodes	Current Node	Objective Bounds	Work					
Expl Unexpl	Obj Depth IntInf	Incumbent BestBd Gap	It/Node Time					
0	0	0.24802	0.6296	6.000000	0.24802	95.9%	-	159s
H	0	0	0	1.0000000	0.24802	75.2%	-	159s
0	0	0.24802	0.6296	1.00000	0.24802	75.2%	-	159s

Explored 1 nodes (1513446 simplex iterations) in 159.98 seconds (83.79 work units)

Thread count was 4 (of 4 available processors)

Solution count 2: 1 6

Optimal solution found (tolerance 1.00e-04)

Best objective 1.000000000000e+00, best bound 1.000000000000e+00, gap 0.0000%

Optimal solution found.

Minimum number of workers for: 97.725172413793135% completion: 22253

Ausführungszeit: 372.38012210721276 min

Set parameter TimeLimit to value 6000

Gurobi Optimizer version 10.0.1 build v10.0.1.rc0 (win64)

CPU model: Intel(R) Xeon(R) Platinum 8171M CPU @ 2.60GHz, instruction set [SSE2]AVX[AVX2]AVXS12
Thread count: 4 physical cores, 4 logical processors, using up to 4 threads

Academic license - for non-commercial use only - registered to nlls.hellmann@stud-mail.uni-wuerzburg.de
Optimize a model with 6298 rows, 1512006 columns and 3038048 nonzeros

Model fingerprint: 0x4d47dda5

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Coefficient statistics:

Matrix range [1e+00, 1e+00]

Objective range [1e+00, 1e+00]

Bounds range [1e+00, 1e+00]

RHS range [1e+00, 1e+00]

Found heuristic solution: objective 6.00000000

Presolve removed 0 rows and 0 columns (presolve time = 5s) ...

Presolve time: 6.49s

Presolved: 6298 rows, 1512006 columns, 3038048 nonzeros

Variable types: 0 continuous, 1512006 integer (1512006 binary)

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
0	2.4801587e-01	0.000000e+00	0.000000e+00	10s

Use crossover to convert LP symmetric solution to basic solution...

Root crossover log...

1510998	PPushes remaining with Pinf 0.0000000e+00	10s
1475939	PPushes remaining with Pinf 0.0000000e+00	10s
800904	PPushes remaining with Pinf 0.0000000e+00	15s
227748	PPushes remaining with Pinf 0.0000000e+00	20s
0	PPushes remaining with Pinf 0.0000000e+00	22s
0	DPushes remaining with Dinf 0.0000000e+00	22s

Push phase complete: Pinf 0.0000000e+00, Dinf 4.3723705e-15

Root simplex log...

Iteration	Objective	Primal Inf.	Dual Inf.	Time
1511001	2.4801587e-01	0.000000e+00	0.000000e+00	22s

Root relaxation: objective 2.480159e-01, 1511001 iterations, 14.46 seconds (11.20 work units)

Nodes	Current Node	Objective Bounds	Work			
Expl Unexpl	Obj Depth IntInf	Incumbent BestBd Gap	It/Node Time			
0	0	0.24802	0.6296	6.000000	0.24802	95