

## Расширения для ноутбуков Jupyter

🕒 Дата публикации Dec 7, 2018

Jupyter Notebook - это отличное обучение, изучение и грамотность [среда программирования](#) Но в готовых ноутбуках, как известно, не хватает функций. К счастью, есть [количество способов](#), в том числе [Расширения Jupyter Notebook](#), чтобы улучшить этот бесценный инструмент.

## Очень сжатая версия

Запустите следующее в командной строке:

```
pip install jupyter_contrib_nbextensions && jupyter contrib nbextension install
```

Запустите Jupyter Notebook и перейдите на новую вкладку Nbextensions:

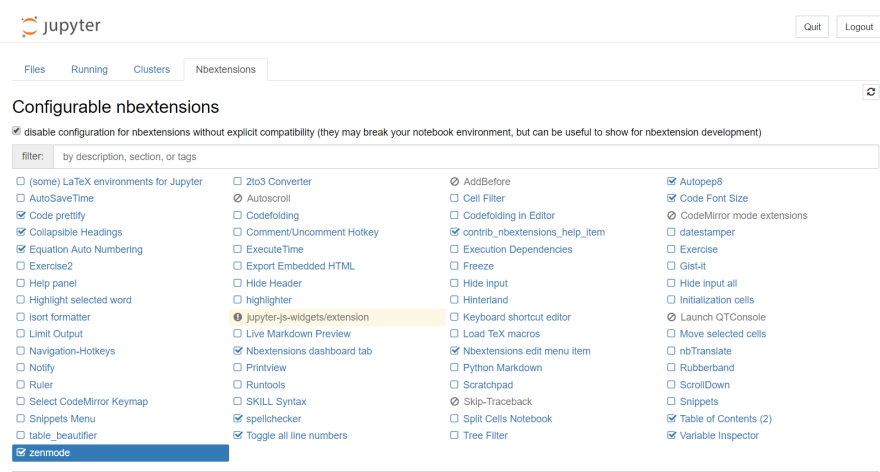


Вкладка «Расширения ноутбука Jupyter»

Включите нужные вам расширения и наслаждайтесь преимуществами производительности.

(Если ты [не вижу вкладку](#) откройте записную книжку и нажмите «Правка»> «Конфигурация nbextensions»)

Включенные расширения можно увидеть на панели инструментов в ноутбуке:



# Чуть длиннее версия

Если вам этого недостаточно, ниже приведены некоторые подробности о расширениях ноутбуков Jupyter. Я также включил мою пятерку лучших, чтобы вы начали.

## Что такое расширения для ноутбуков?

[Расширения Jupyter Notebook](#)это простые надстройки, которые расширяют базовую функциональность среды ноутбука. Написанные на JavaScript, они делают такие вещи, как[Автоформат вашего кода](#)или[отправить уведомление в браузер](#)когда ячейка завершена. Расширения в настоящее время работают только в Jupyter Notebooks (не в Jupyter Lab).

Зачем использовать эти расширения? Ноутбуки Jupyter являются отличными инструментами для преподавания, изучения, создания прототипов, изучения и опробования новых методов (или даже [производство на Netflix](#)). Тем не менее, ванильные ноутбуки ограничены в функциях, которые могут расстроить работу в них. Хотя расширения Jupyter Notebook не решают проблему полностью, они добавляют несколько преимуществ, которые облегчат вашу работу.

## Какой использовать

Ниже приведены 5 расширений Jupyter Notebook, которые я использую чаще всего:

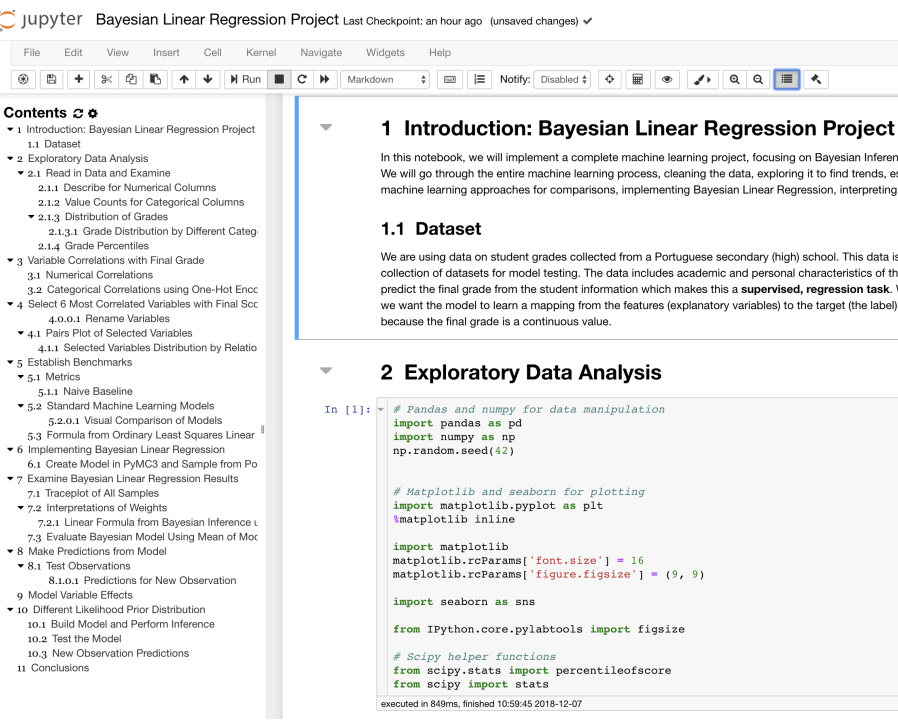
### 1. Содержание: удобная навигация

Как только вы начнете получать десятки ячеек в одной записной книжке Jupyter, может быть трудно отслеживать их все. Оглавление решает эту проблему путем добавления связанного содержания, которое можно разместить в любом месте на странице:



#### Оглавление

Вы также можете использовать расширение, чтобы добавить связанное оглавление вверху блокнота. Это даже показывает, какая ячейка выбрана, а какая запущена:



#### Содержание в тетради

### 2. Autoper8: аккуратный код в один клик

Мы все должны написать[совместимый с пер8 код](#)Но иногда вы попадаете в анализ, и трудно придерживаться лучших практик. Когда вы закончите писать этот удивительный сюжет, это расширение позволит вам просто щелкнуть молотком и автоматически отформатировать ваш грязный код.

Table of Contents

- 1 Introduction: Bayesian Linear Regression Project
  - 1.1 Dataset
- 2 Exploratory Data Analysis
  - 2.1 Read in Data and Examine
    - 2.1.1 Describe for Numerical Columns
    - 2.1.2 Value Counts for Categorical Columns
    - 2.1.3 Distribution of Grades
      - 2.1.3.1 Grade Distribution by Different Categorical Variables
    - 2.1.4 Grade Percentiles
- 3 Variable Correlations with Final Grade
  - 3.1 Numerical Correlations
  - 3.2 Categorical Correlations using One-Hot Encoding
- 4 Select 6 Most Correlated Variables with Final Score
  - 4.0.0.1 Rename Variables
  - 4.1 Pairs Plot of Selected Variables
    - 4.1.1 Selected Variables Distribution by Relation to Median
- 5 Establish Benchmarks
  - 5.1 Metrics
    - 5.1.1 Naive Baseline
  - 5.2 Standard Machine Learning Models
    - 5.2.0.1 Visual Comparison of Models
  - 5.3 Formula from Ordinary Least Squares Linear Regression
- 6 Implementing Bayesian Linear Regression
  - 6.1 Linear Model with Priors and Gaussian Process Kernel
- 7 Examine Bayesian Linear Regression Results
  - 7.1 Traceplot of All Samples
  - 7.2 Interpretations of Weights
    - 7.2.1 Linear Formula from Bayesian Inference using Mean of Parameters
  - 7.3 Evaluate Bayesian Model Using Mean of Model Parameters
- 8 Make Predictions from Model
  - 8.1 Test Observations
    - 8.1.0.1 Predictions for New Observation
- 9 Model Variable Effects
- 10 Different Likelihood Prior Distribution
  - 10.1 Build Model and Perform Inference
  - 10.2 Test the Model
  - 10.3 New Observation Predictions
- 11 Conclusions

1 Introduction: Bayesian Linear Regression Project

In this notebook, we will implement a complete machine learning project, focusing on Bayesian Inference methods, in particular, Bayesian Linear Regression. We will go through the entire machine learning process, cleaning the data, exploring it to find trends, establishing a baseline model, evaluating several machine learning approaches for comparisons, implementing Bayesian Linear Regression, interpreting the results, and presenting the results. Let's get started!

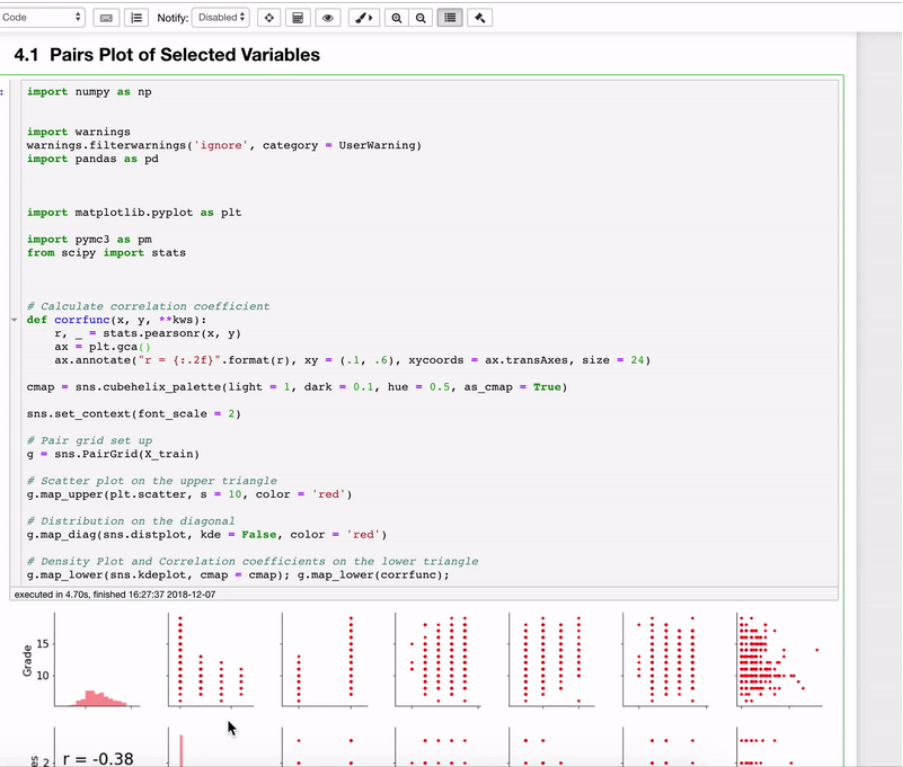
1.1 Dataset

Автор пер8 ваш код.

Как и лучшие надстройки, эта задача выполняет трудоемкую и утомительную задачу простым щелчком мыши, позволяя вам сосредоточиться на решении проблем.

3. Инспектор переменных: следите за своим рабочим пространством

Инспектор переменных показывает имена всех переменных, которые вы создали в записной книжке, а также их тип, размер, форму и значение.



Переменный инспектор

Этот инструмент неоценим для специалистов по обработке данных, мигрирующих из RStudio, или если вы не хотите продолжать печатать `df.shape` или не могу вспомнить `type` из `x`,

4. ExecuteTime: показать, когда и как долго работали клетки

Я часто ловлю себя на том, что пытаюсь выяснить, сколько времени понадобилось для работы камеры или когда я последний раз запускал ноутбук, который был открыт несколько дней. ExecuteTime позаботится об обоих из них, показывая, когда ячейка закончила и сколько времени это заняло.

Variable Inspector					
X	Name	Type	Size	Shape	Value
x	ElasticNet	ABCMeta	1464		
x	ExtraTreesReg...	ABCMeta	2000		
x	GradientBoost...	ABCMeta	2000		
x	LinearRegress...	ABCMeta	1184		
x	RandomForestR...	ABCMeta	2000		
x	SVR	ABCMeta	2000		
x	X_plot	DataFrame	27018	(474, 7)	failures higher_edu mother_edu...
x	X_test	DataFrame	15455	(159, 8)	Grade failures higher_edu mot...
x	X_train	DataFrame	47498	(474, 7)	Grade failures higher_edu mot...
x	all_model_res...	DataFrame	512	(8, 2)	mae rm...
x	ax	AxesSubplot	56		AxesSubplot(0.569606,0.345104,0.40634...
x	category_df	DataFrame	111632	(633, 17)	school sex address famsize Pstatu...
x	cmap	ListedColormap	56		
x	col	str	93		absences
x	df	DataFrame	197720	(633, 34)	school sex age address famsize P...
x	dummy_df	DataFrame	57827	(633, 44)	school_GP school_MS sex_F sex...
x	family	Normal	56		Family
x	formula	str	127		Grade ~ failures + higher_edu + mothe...
x	g	PairGrid	56		
x	i	int	28		5
x	lr	LinearRegress...	56		LinearRegression(copy_X=True, fit_int...
x	mb_mae	float64	32		2.1761006289308176
x	mb_mse	float64	32		2.6776503357897044
x	median_pred	float	24		12.0
x	median_preds	list	1448		[12.0, 12.0, 12.0, 12.0, 12.0, 12.0, ...]
x	model_formula	str	223		Grade = 9.18 * Intercept + -1.31 * f...
x	normal_model	Model	56		
x	normal_trace	MultiTrace	56		
x	observation	Series	56	(7,)	Intercept 1 absences 1 failu...
x	ols_formula	str	180		Grade = 9.19 + -1.32 * failures + 1.8...
x	results	DataFrame	512	(8, 2)	mae rm...
x	schools	Series	32	(4,)	school address GP U 34...
x	subset_above	DataFrame	14934	(262, 7)	failures higher_edu mother_edu...
x	subset_below	DataFrame	12084	(212, 7)	failures higher_edu mother_edu...

Выход расширения ExecuteTime

Есть лучше[методы для выбора времени](#)(такие как[%%timeit](#)[клеточная магия](#)), но это легко реализовать и охватывает все ячейки в ноутбуке.

## 5. Скрыть ввод кода: скрыть работу, показать результаты

Хотя некоторым из нас нравится видеть тяжелую работу, которая входит в анализ, некоторым людям просто нравится видеть результаты. Расширение Hide input all позволяет мгновенно скрывать весь код в блокноте, сохраняя результаты.

```
# Context for model
with pm.Model() as t_model:
    # Family is Student's T in this case
    family = pm.glm.families.StudentT(df = 2)

    # Formula, data, family
    pm.GLM.from_formula(formula, data = X_train, family = family)

    # Sample from the posterior
    t_trace = pm.sample(draws=2000, tune=500, njobs = -1)
```

executed in 52m 16s, finished 15:47:11 2018-12-07

Скрыть весь код

В следующий раз, когда кто-то скажет, что он просто хочет увидеть результаты, у вас есть решение для одного клика. (Хотя вы всегда должны проверять код).

Это только те расширения, которые я использую чаще всего, и вы не ограничены пятью. Проверьте весь список, установив расширения и открыв блокнот (блокнот показан[вот на GitHub](#))! (Если вы готовы, вы даже можете написать свое собственное расширение. Документация довольно скудная, но[вот основная структура](#).)

## Выводы

Установите расширения Jupyter Notebook, потратьте некоторое время на выяснение, какие из них полезны для вас, и повысьте свою производительность. Хотя ни один из них не меняет жизнь, все они приносят достаточную выгоду, чтобы быть полезными, в совокупности экономя вам часы драгоценного времени на разработку.

Хотя вы, вероятно, захотите потратить некоторое время на[изучение IDE](#)если вы пишете производственный код (мне нравится[Кодекс VS](#)), Ноутбуки Jupyter по-прежнему являются неотъемлемой частью процесса обработки данных Если вы собираетесь использовать эту среду, вы также можете получить максимум от своих инструментов.

Как всегда, я приветствую отзывы и конструктивную критику. Со мной можно связаться в твиттере[@koehrsen will](#)или через мой личный сайт[willk.online](#),

- [Фреймворки и библиотеки \(большая подборка ссылок для разных языков программирования\)](#)
- [Список бесплатных книг по машинному обучению, доступных для скачивания](#)
- [Список блогов и информационных бюллетеней по науке о данных и машинному обучению](#)
- [Список \(в основном\) бесплатных курсов машинного обучения, доступных в Интернете](#)