Week – 1
ML &
Python 101

ML Bootcamp 2021

GDG Cloud EDMONTON

AlbertaAI

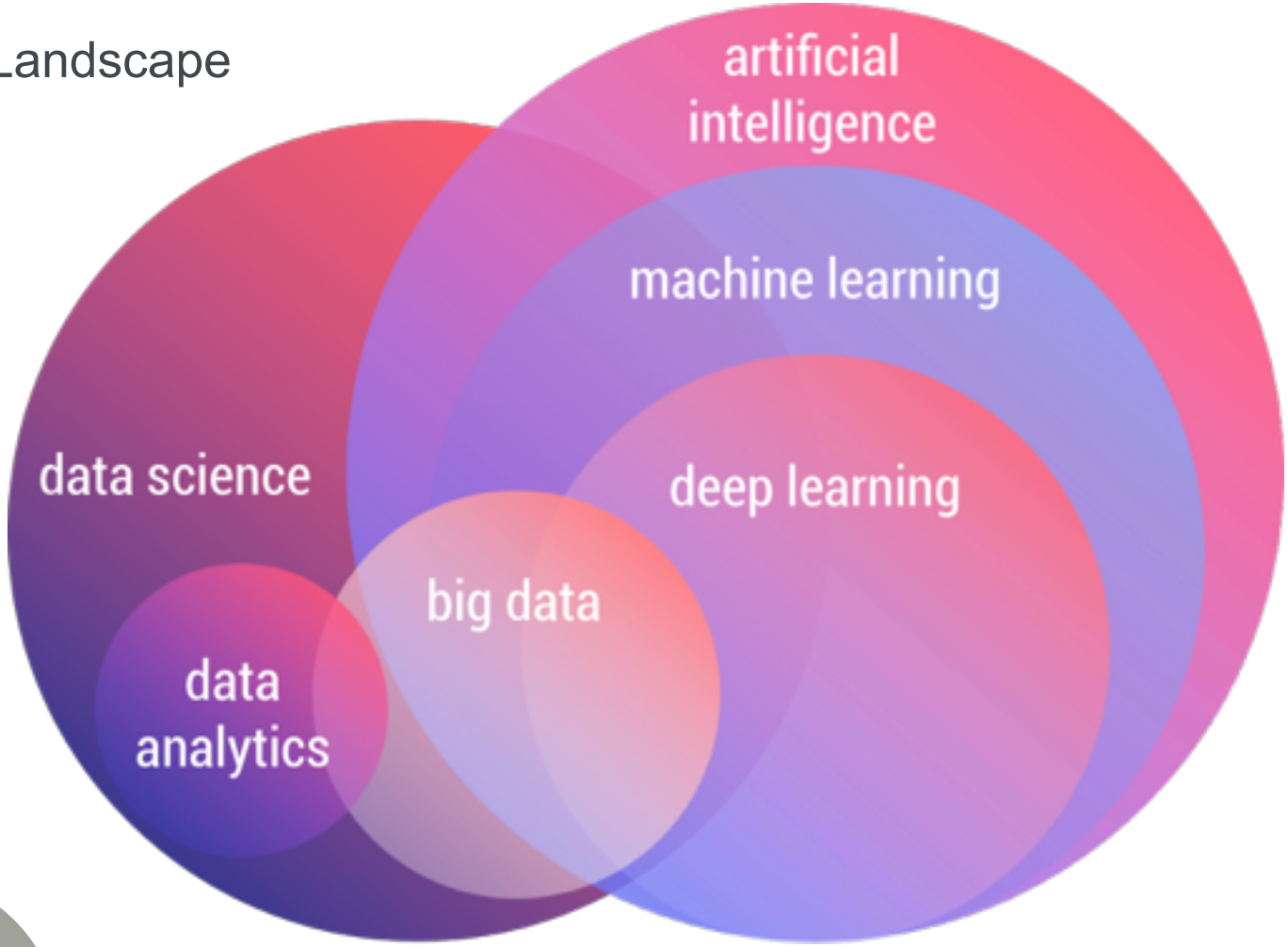
Careera Analytics Lab

Slido: #ml_bootcamp_week1

Table of contents

Slido: #ml_bootcamp_week1

The AI Landscape

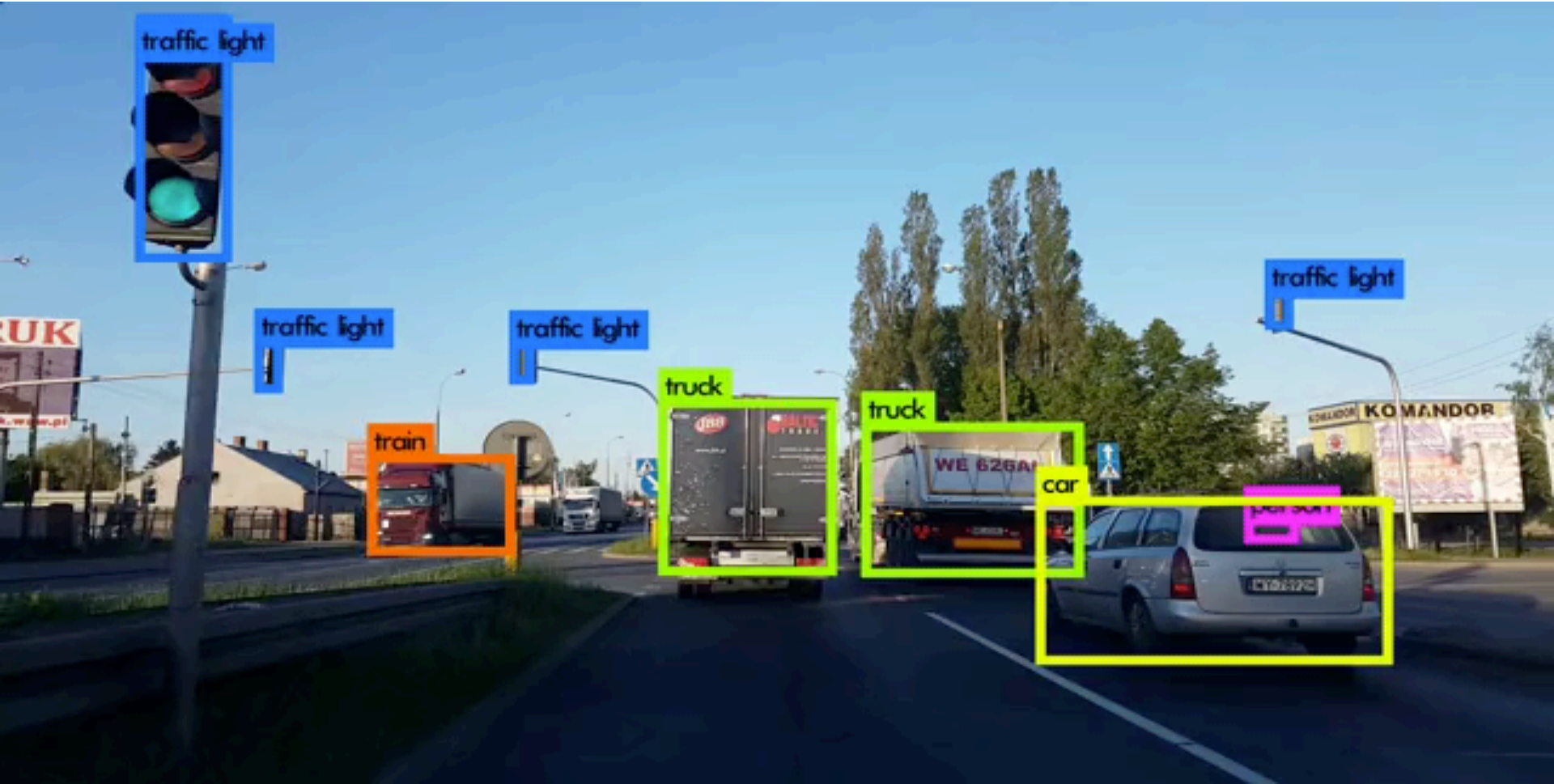


Applications of Artificial Intelligence

- **Computer Vision**: Face recognition, Object detection, Video Quality Assessment
- **Natural Language Processing**: Next word prediction(email, phone etc.), Question-Answering, Text Summarization
- **Time Series Forecasting**: Stock market trend predictions, weather predictions
- **Fraud Detection**: Detecting Outliers/Anomalies in banking and other related fields

Slido: #ml_bootcamp_week1

Computer Vision in action



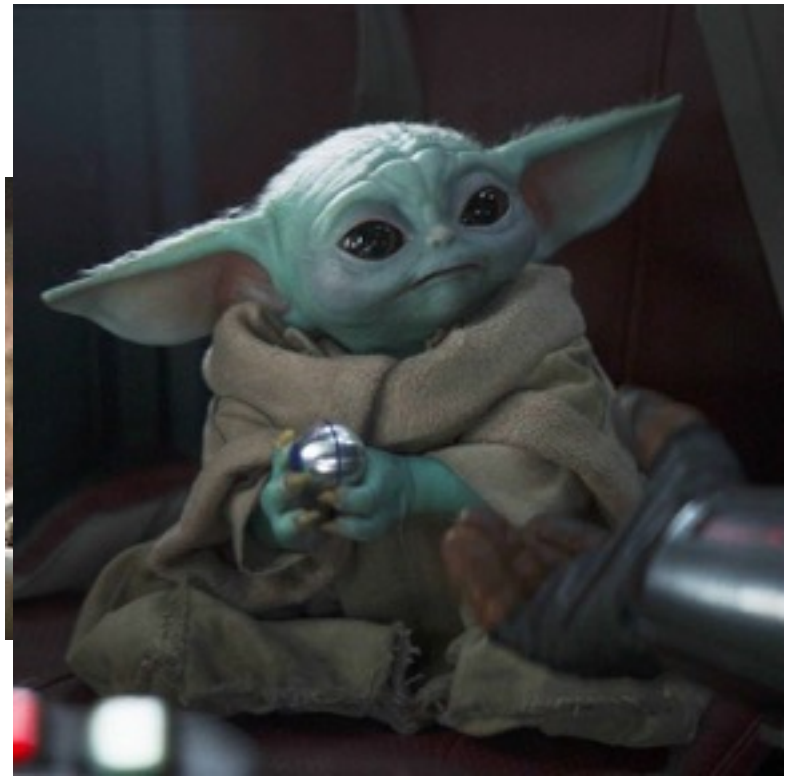
“Reel” AI

“Real” AI

```
# define workflow
class TaskGetData(d6tflow.tasks.TaskPqPandas):  # save dataframe as parquet

    def run(self):
        iris = sklearn.datasets.load_iris()
        df_train = pd.DataFrame(iris.data,columns=['feature{}'.format(i) for i in range(4)])
        df_train['y'] = iris.target
        self.save(df_train) # quickly save dataframe

class TaskPreprocess(d6tflow.tasks.TaskCachePandas):  # save data in memory
    do_preprocess = luigi.BoolParameter(default=True) # parameter for preprocessing yes/no

    def requires(self):
        return TaskGetData() # define dependency
```

Slido: #ml_bootcamp_week1

The Yoda Challenge



BABY YODA

The ML Tech Stack

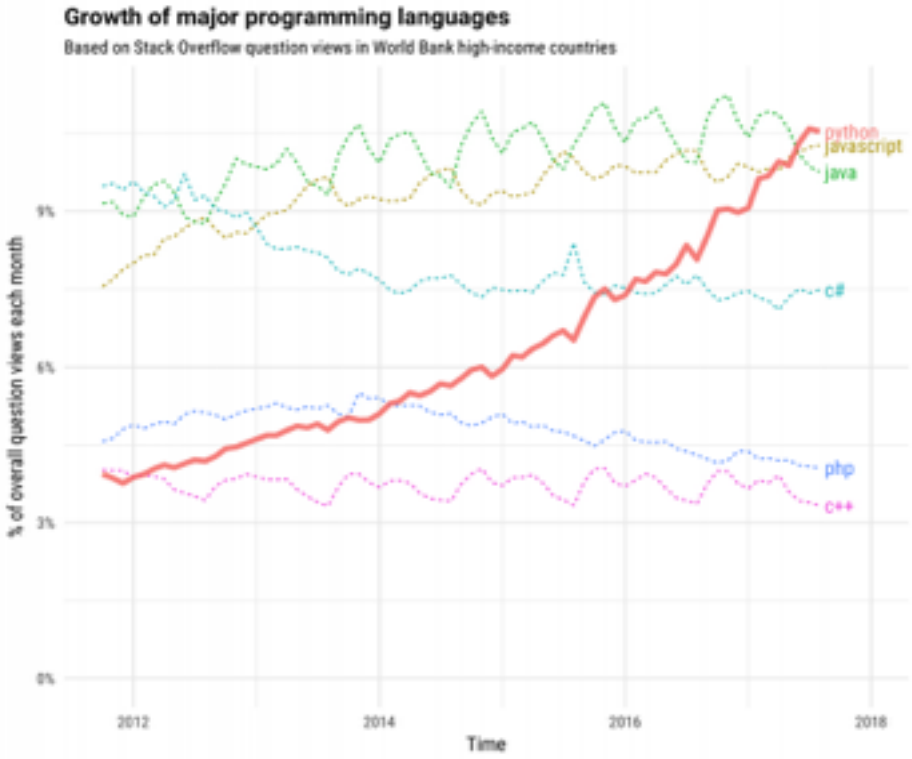
- Good Programming skills
- Python
- Machine Learning Algorithms
- Basic Statistics & Probability concepts
- Patience to read research papers
- Databases and SQL
- Application Development concepts: Web apps, cloud computing
- Some idea behind Data Lakes and related concepts

Our focus for today

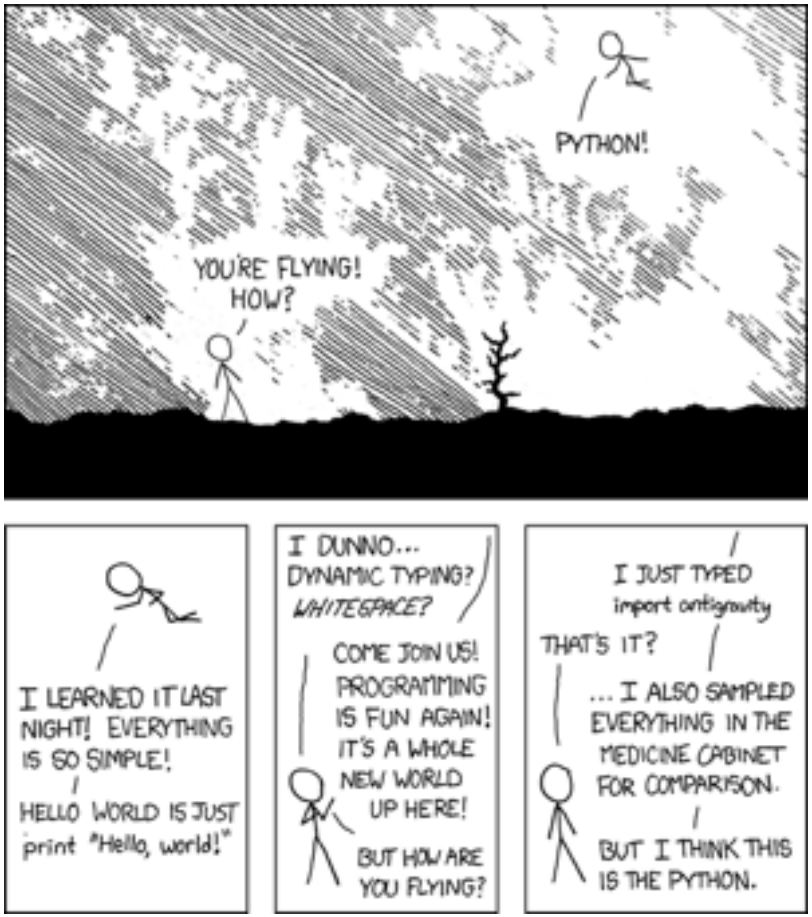
Slido: #ml_bootcamp_week1

Why is Python so popular?

- Human-like language
- Easy to read and understand
- Interpreted language
- Problem solution first, implementation later
- Libraries are easy to use



Slido: #ml_bootcamp_week1


PYTHON!
YOU'RE FLYING! HOW?
I LEARNED IT LAST NIGHT! EVERYTHING IS SO SIMPLE!
HELLO WORLD IS JUST print "Hello, world!"
I DUNNO... DYNAMIC TYPING? WHITESPACE?
COME JOIN US! PROGRAMMING IS FUN AGAIN! IT'S A WHOLE NEW WORLD UP HERE!
BUT HOW ARE YOU FLYING?
I JUST TYPED import antigravity
THAT'S IT?
... I ALSO SAMPLED EVERYTHING IN THE MEDICINE CABINET FOR COMPARISON.
BUT I THINK THIS IS THE PYTHON.

ML jargon

- **Datasets:** The data that you process, transform and use to train your ML models
- **Features:** The individual columns(if it is a table) in your data
- **Supervised/Unsupervised:** Either with or without answers
  - **Classification:** Predicting a limited set of outputs Eg: Pos/Neg sentiment, Cat/Dog, Bearish/Bullish
  - **Regression:** Output has an infinite possibilities Eg: Housing prices, stock prices, time etc.
- **Machine Learning:** A set of algorithms or a method to learn through experience using data
- **Deep Learning:** Just a fancy word for Neural Networks
- **Training and Testing:** We usually split our data into two parts, one to train the model and another to test the model's performance
- **Functions:** A block of code with a name that can be called to execute a set of statements
- **Library:** A collection of functions grouped under one single name
- **Package:** A collection of script (*.py) files containing modularized code

Slido: #ml_bootcamp_week1

Commonly used libraries

Pandas

Popular library for data transformations and manipulations

Numpy

A scientific computing package for Python

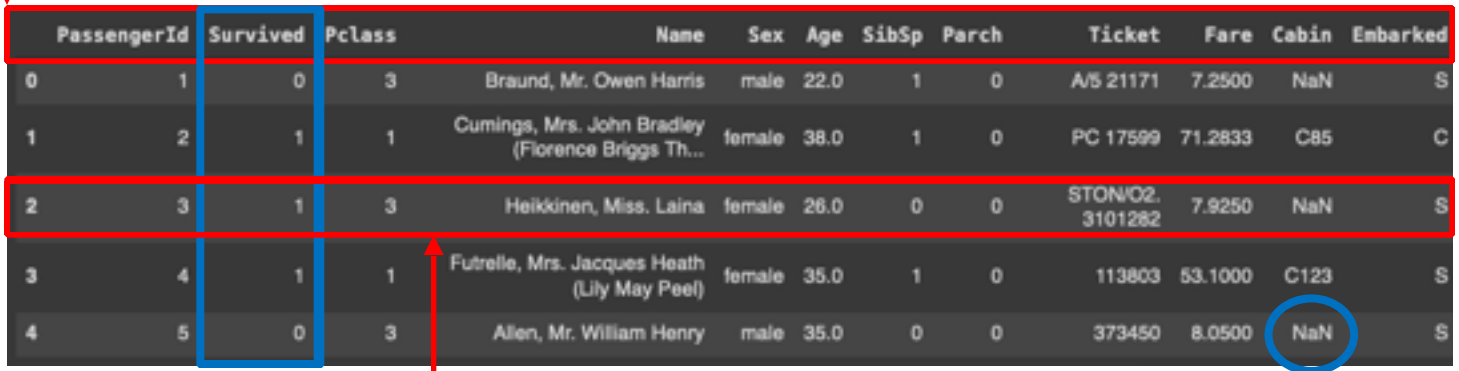
Scikit-learn

The home of all Machine Learning Algorithms!

Slido: #ml_bootcamp_week1

Features/Columns

Pandas Dataframe



	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

Pandas Series object

Data point

String

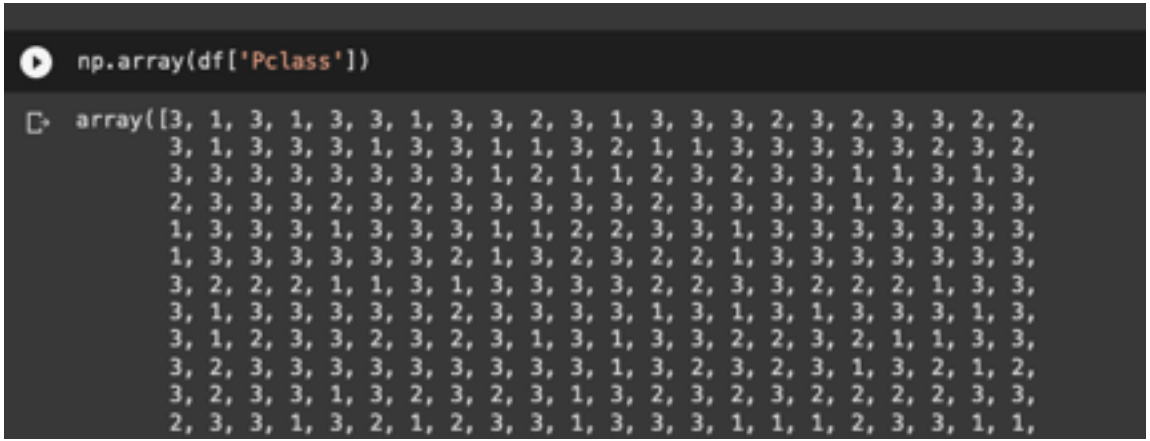
Categorical

Numerical

Missing Value

Numpy

- At the core of the NumPy package, is the *ndarray* object -- *numpy.org*



Lets start Practicing!

Slido: #ml_bootcamp_week1

Thanks

Do you have any questions?


slidesgo