

## 1、 演讲主题；

经济学及金融学实证研究中的大数据分析：基于 Python 和 PostgreSQL

## 2、 演讲主题简介 ( 300 - 600 字 )；

### 1).概要介绍：

经济学和金融学中的实证研究在技术上依赖于对计量方法的掌握，在主题上则依赖于对数据的获取和处理能力。而随着宏观经济数据、工业企业数据库、海关数据库、人口普查与抽样调查数据、农村金融/家庭金融调查数据、低频股价数据、上市公司财务及治理数据的广泛使用，其中蕴含的高价值研究主题已日渐稀缺，转向数量更为庞大但处理也更为艰难的数据成为必然。

而这些庞大的新数据集，无论工商数据、裁判文书数据、股票论坛数据、高频股价数据还是建立在详细地理坐标下的各类时空数据，均庞大无比，甚至于庞大之外还添杂乱：动辄百 G，甚至数 T；散布于多张表，甚至于文档之中。传统的以数据分析为目标的 SAS、SPSS、Stata、Matlab 等软件，碰上这类数据，即便不至于束手无策，常常也难比登天。

本讲座将介绍一个 Python 语言与 PostgreSQL 数据库相结合的并行大数据处理方案，该方案专为这类大规模数据处理与分析而设，将助你轻松遨游大数据的海洋，在更广阔的主题上开展激动人心的经济学及金融学实证研究。

### 2).内容大纲：

#### 1. “经济学-金融学” 研究中的普遍规律

- 1.1 科学研究的基础
- 1.2 研究对象的分佈
- 1.3 实证研究的抽象阶梯
- 1.4 实证研究的生态系统

#### 2. “经济学-金融学” 研究中的大数据化

- 2.1 经济学研究中的大数据化
- 2.2 金融学研究中的大数据化
- 2.3 经济学、金融学大数据的特点

#### 3. “经济学-金融学” 研究中的大数据解决方案

- 3.1 整体方案：Python + PostgreSQL
- 3.2 为何选 Python
- 3.3 为何选 PostgreSQL
- 3.4 实证分析中的 PostgreSQL 配置方案

#### 4. “经济学-金融学” 中的大数据并行方案

- 4.1 单机版并行方案
- 4.2 集群版并行方案

#### 4.3 应用场景举例

**3、 演讲适合听众范围（即用户需要什么样技术基础，才能更好地交流和学习，并从中获益）；**  
无需任何基础，只需对“编程与大数据分析在经济学等社会科学”中的应用状况感兴趣即可。