New College of Florida, January 18-22, 2021, 8:30AM – 12:30PM

Presenters
> Bahar Fardanian
> Lili Xu
> Arjuna Chala
> Dan Camper

Overview
> HPCC Systems is a distributed big data technology used by LexisNexis Risk Solutions. It was, in fact, created by LNRS long before Hadoop came along and it still serves as the core technology behind a $3B/year business.
>
> In this workshop you will be introduced to the HPCC Systems technology and to the data flow language used to program on it, ECL. The data you will be examining is a detailed global schedule of airline flights. Each record contains not only takeoff/landing places and times, but details regarding the plane, meals, and crew are included as well.

Agenda
> **All days**
>> Four-hour workshop
>> There will be at least one 10-15 minute break each day
>
> **Tuesday, January 19**
>> Introduction to HPCC Systems
>> Introduction to ECL
>> Student CloudIDE setup
>> Data Discovery — getting acquainted with the flight data
>> Review data dictionary
>> Profiling the actual data
>> Manual inspection via simple ECL
>> ECL concepts
>> Practical ECL with flight data
>> Arjuna Chala: The New Big Data Journey
>
> **Wednesday, January 20**
>> ECL concepts (cont'd)
>> Practical ECL with flight data (cont'd)
>
> **Thursday, January 21**
>> ECL (cont'd)
>> Practical ECL with flight data
>> Dan Camper: Data Profiling

Friday, January 22
ECL Machine Learning
Practical ECL ML

Student Requirements
Each student should bring a laptop

Student Pre-work (recommended, but optional)
Learn about ECL
HPCC Systems ECL Concepts videos
[HPCC Systems ECL Concepts Part 1](#)
[HPCC Systems ECL Concepts Part 2](#)
[HPCC Systems ECL Concepts Part 3](#)
[HPCC Systems ECL Concepts Part 4](#)
[HPCC Systems ECL Concepts Part 5](#)
[HPCC Systems ECL Concepts Part 6](#)
Learn ECL: https://hpccsystems-solutions-lab.github.io/