MOOCでのcertification問題について

櫻井尚子・東京情報大学
和泉志津恵・大分大学
深澤弘美・東京医療保健大学

agenda

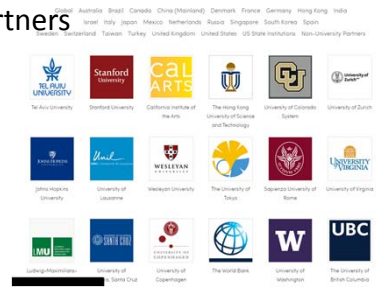
1. MOOC の動向
2. 統計関係授業の紹介
3. 修了認定の中身
4. 評価方法の例
5. まとめ

MOOCの概略

Massive Open Online Courses
2012年前半～
大学⇔資本⇔ネットワーク技術
例： Coursera
edX
UDACITY：Sebastian Thrun with Google
「AI入門」のオンラインコース（2011）

coursera

【2015年2月25日時点】
Join 11,655,477 Courserians
Learn from **965 courses**
from our 116 partners



Statistics & Data Analysis

73 コース → (全言語)

コンピュータサイエンス関連→211
数学→76
ソーシャルサイエンス→163
人文科学→175

edX → Data Analysis, Science

Statistics & Data Analysis に45 コース

- **The Analytics Edge (MIT)**
  R を使って，データから現実世界へ
- **Introduction to Big Data with Apache Spark (UC Berkeley)**
  「ポストHadoop」
  計算処理を分散環境にて並列実行するための基盤
- **Statistics and R for the Life Sciences (Harvard)**
  Distributions, Exploratory Data Analysis, Inference, Non-parametric statistics

....

Udacity → Data Science

16 コース
Learn data science from **industry experts** at Facebook, Cloudera, MongoDB, Georgia Tech, and more. 実際の企業のプロから実習

Cloudera：Hadoopのプラットフォーム
Hadoop（Apache **Hadoop**）（Apache Software Foundation(ASF)が開発・公開している，大規模データを効率的に分散処理・管理するためのソフトウェア基盤(ミドルウェア)．オープンソースソフトウェアとして公開）
MongoDB：
ドキュメント指向DB（オープンソース），英語版）．NoSQL.

Data Analyst Nanodegree

Discover Insights from Data

データがもたらす意味を発掘しよう

You will work with your **peers and advisors** on projects approved **by leading employers** as the critical indicators of job-readiness.

仕事の現場と直結

We designed these projects with **expert** Data Analysts, Data Scientists, and **hiring managers**.

データサイエンティストの不足

The Data Scientist's Tool Box (Coursera)

overview of the **data, questions, and tools** that data analysts and data scientists work with.

2 components :

1. the ideas behind turning data into **actionable knowledge** **現実的な知識へ**
2. Practical **tools** : version control, markdown, git, Github, R, Rstudio **実現への道具類**

Course content 中身

- Track motivation
- Getting help
- Introduction to basic tools R
- Rstudio
- Git : ﾊﾞｰｼﾞｮﾝ管理ｼｽﾃﾑ
- Github :ｿﾌﾄｳｪｱ開発ﾌﾟﾛｼﾞｪｸﾄのための共有ｳｪﾌﾞｻｰﾋﾞｽ
- Types of data questions:ﾃﾞｰﾀがもたらす様々な問題
- Steps in a data analysis
- Putting the science in data science

Course 材料

ビデオレクチャー

Weekly quizzes（3回までトライ可）

late days（5回までは評価に影響なし）

soft deadlines & hard deadlines

＋

The Course Project

締切 deadline

各quiz の締切 => soft deadline

それより5日後 => hard deadline

late day => さらなる遅延

（10% down）

※各自 late days を5回まで使用可能.

（これは評価に影響しない）

コースの流れ The Data Scientist's Tool Box

Announcements
Welcome to the Data Scientist's Toolbox!
Pre-Course Survey（ｱﾝｹｰﾄ調査） いろいろなﾊﾞｯｸｸﾞﾗｳﾝﾄﾞ
Course Project: Setting Up Accounts
the basic software setup
Data Science Specialization Community Site（ｽﾍﾟｼｬﾙの1つ）
Week 2 : Installing the Toolbox 道具類の導入
Week 3 : Conceptual Issues（ﾃﾞｰﾀ解説）
Course Project Submissions Due on Sunday Evening（ﾌﾟﾛｼﾞｪｸﾄ提出）
Week 4 : Practice with the Tools and Peer Review（相互評価）
Data Scientist's Toolbox: Course Project Evaluations Due on Sunday Evening（最終評価）

Community Site

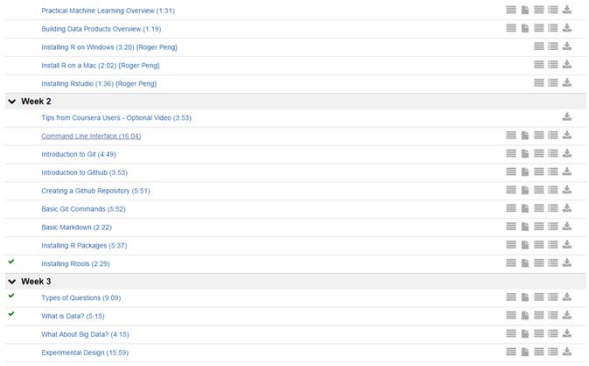
Unbelievable passion students の存在
2名の素晴らしい生徒→ディスカッションを通じて，ハイクォリティなワークの提出

GitHub を使ってコミュニティ作成，生徒たちが自由に投稿→レベルの高いワークにいち早く出会える→全員で共有できる→Data Science Specialization 全コースに向けて提供できる

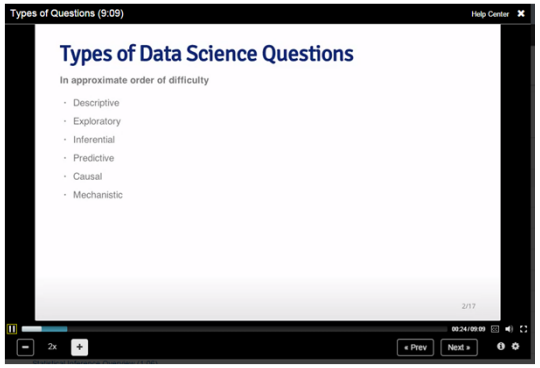
Conceptual Issues（第3週）
ﾃﾞｰﾀについて

Week 3
データを知識と化す
疑問点はﾒｯｾｰｼﾞﾎﾞｰﾄﾞ上で探して下さい
既にｽｰﾊﾟｰｽﾀｰの人は，他の人の疑問に答えてほしい（手伝いをして下さい）→このことが，君の能力を上げ，人に教える力を増幅させる．さらには，優秀なﾃﾞｰﾀｻｲｴﾝﾃｨｽﾄへの道につながる．

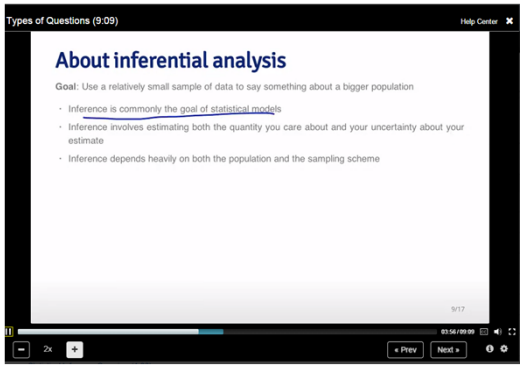
Week 3 video list



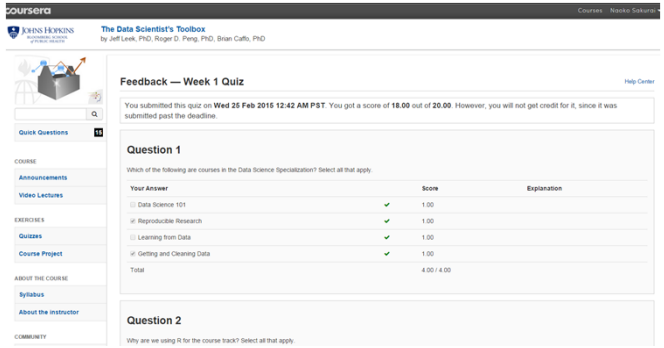
ビデオ例



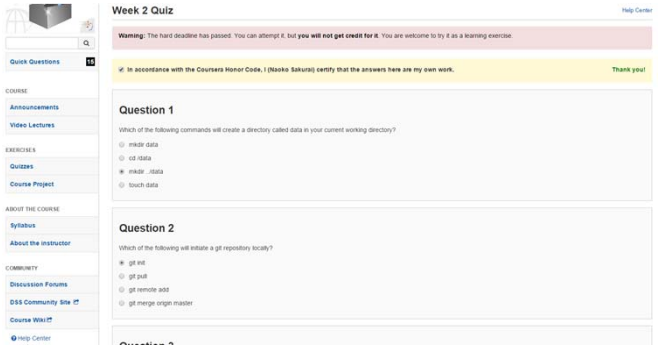
inferential



Week 1 Quiz



Week 2 Quiz



Week 3 Quiz

シラバスから

剽窃（ひょうせつ）（盗作等）→ Johns Hopkins University の規定を適用

I will register for only one account. My answers to homework, quizzes and exams will be **my own work** (except for assignments that explicitly permit collaboration). I will not make solutions to homework, quizzes or exams available to anyone else. This includes both solutions **written by me**, as well as any official solutions provided by the course staff. I will not engage in any other activities that will dishonestly improve my results or dishonestly improve/hurt the results of others.

Grading policy

Quiz 1 = 20%

Quiz 2 = 20%

Quiz 3 = 20%

Course project = 40%

合格ライン

You must receive a final grade of **70% or better** to pass the course and achieve the certificate. You must receive a final grade of **90%** to pass the course **with distinction**.

スペシャル認定

※相互評価を実施しないと，自身の成績は 20% ダウン

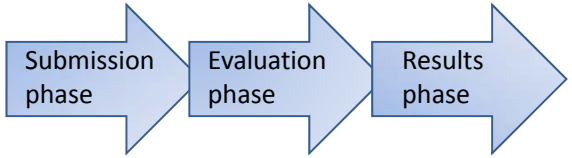
Course project

ソフトウェアのインストールを行い、その evidence をファイルやテキストにて提出する

R と Rstudio を統計計算の道具として用い，cloud ソフトウェアを提出物の共有に利用

インストールの方法はシンプル

説明内容を理解する英語力が必要



Course project original (1)

① Install R
② Install Rstudio
} R/RStudioのインストール

③ Open Rstudio and take a screenshot
④ Submit a **screenshot** of Rstudio open on your screen using one of these formats: png, jpg, gif, pdf → スクリーンショットを指定場所へ

※ Be careful not to have personally identifiable or important information visible in the screen shot!

個人情報流出への注意

Course project original (2)

① Set up a **Github** account (you may use your own name or a pseudonym).
Github アカウントをつくりなさい.

② Create a repo called datasciencecoursera
datasciencecoursera という名称の格納庫を作成しなさい.

③ Submit the link to your GitHub account (or a direct link to your datasciencecoursera repo)
そこへリンクをはりなさい.

Course project original (3)

① Create a text file called HelloWorld.md
ﾏｰｸﾀﾞｳﾝ: 自動的にHTMLコードとして変換して出力

② Add the line "## This is a markdown file" (without the quotation marks) to the document (without the quotation marks)
ラインの追加

③ Push the document to the datasciencecoursera repo you created on Github
上記ドキュメントをGithub格納庫へ

④ Submit the link to the HelloWorld.md file on your Github repo.
リンクを実行

Course project original (4)

① Fork the data sharing repository here: https://github.com/jtleek/datasharing
ﾃﾞｰﾀを指定の保管場所へ移す

② Submit the link to the forked repository on your Github account.
上記場所へリンクをはる

Peer assessment 相互評価

コースプロジェクトは非常に大事．締切は守ること．評価は全体の40%に相当．

Peer Assessments

The Course Project is an important component of this course and most of the other courses in the Data Science Specialization. Course Projects are usually evaluated and graded through peer assessment. You must submit your Course Project by the specified deadline to receive credit for your submission. Course Project deadlines are not flexible, and you may not apply Late Days to Course Projects.

The Course Project in The Data Scientist's Toolbox is worth 40% of the final grade.

To access the submission interface, scroll to the bottom of this page and click the blue Go To Assignment button.

Time	What happens then
Course Project submission starts at the beginning of Week 1	You can see the Course Project assignment and start working on it. You can save drafts of your work as you go along, and you can come back later to continue working on your draft. When you're ready to submit your work for evaluation, remember to click the "Submit" button. If the deadline passes and you haven't clicked "Submit" yet, then your saved draft will not be evaluated. **Note:** You can submit and re-submit your work for evaluation as many times as you want before the submission deadline, without any penalty. Only your last submission will be seen and evaluated by your classmates.
Submissions due by the end of Week 3 (BEFORE 11:30 PM UTC on Sunday)	After this time, you can no longer change your Course Project submission. If you have not clicked the "Submit" button by this time, your classmates will not see your submission, you will not receive an evaluation for this assignment, and you will not be permitted to evaluate your classmates' submissions.
Evaluation starts at the beginning of Week 4	After this time, you will evaluate the Course Project Submissions of at least four of your classmates. If you don't evaluate at least four, your own grade on the Course Project will be reduced by 20%. Optionally, you can choose to evaluate the work of even more of your classmates before the evaluation deadline passes. This is very helpful for the success of the course. **Remember:** as with the submission stage, this evaluation stage is required if you want your own assignment submission to be evaluated.
Evaluations due by the end of Week 4 (BEFORE 11:30 PM UTC on Sunday)	After this time, you can no longer evaluate the work your peers. If you have not finished your assigned evaluation tasks by this time, your own Course Project Grade will be reduced by 20%.

Assignments

Peer assessment 条件

期日までに，少なくとも4人分のプロジェクトに評価を出す

評価締切を過ぎたら他の人のプロジェクトの結果は見れない

自分への他の人からの評価が送られる

期日までにアセスメント要件を満たさない場合
→自身の評価が20%ダウン

Peer assessment 実際

The Data Scientist's Toolbox | Coursera 2/7 ページ

8人の評価はトップです！

Congratulations on finishing this assessment. To see your results, come back after the evaluation deadline has passed.

You've evaluated 8 submissions! **You're tops!** Want to evaluate another one?

✔ Submitted | Re-submit evaluation

✚ Evaluate another student (optional but useful) (/datascitoolbox-011/human_grading/view/courses/973486/assessments/3/peerGradingSets/5291/next)

Submission from: Student 8

This assignment is designed to make sure you have done the basic software setup that will get you through the rest of the Data Science Specialization. Each component will be evaluated on a yes/no basis by your peers with 10 points assigned for each yes answer.

1. Install R
2. Install Rstudio
3. Open Rstudio and take a screenshot
4. Submit a screenshot of Rstudio open on your screen using one of these formats: png, jpg, gif, pdf

Be careful not to have personally identifiable or important information visible in the screen shot!

Evaluation/feedback on the above work

The link goes to a fork of https://github.com/jtleek/datasharing (https://github.com/jtleek/datasharing) in the user's account.

10: Yes

本人のジョブだと判断できるか？

Overall evaluation/feedback

As far as you can determine, does it appear that the work submitted for this project is the work of the student who submitted it?

1: Yes

あなたの評価結果を述べて下さい．コメントあれば書き添えて下さい．

Please use the space below to provide constructive feedback to the student who submitted the work. Point out the submission's strengths as well as areas in need of improvement. You may also use this space to explain your grading decisions.

Peer assessment 利点

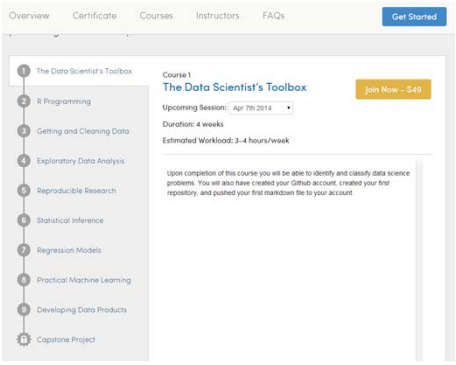
学生全員が評価に参加できる
公平さを保てる**
オープンなので納得を得やすい
アクティブな学習法である
教員の評価負担が少ない
教員は別次元の視点で評価が可能
様々な角度からの評価視点が見える

Peer assessment 課題

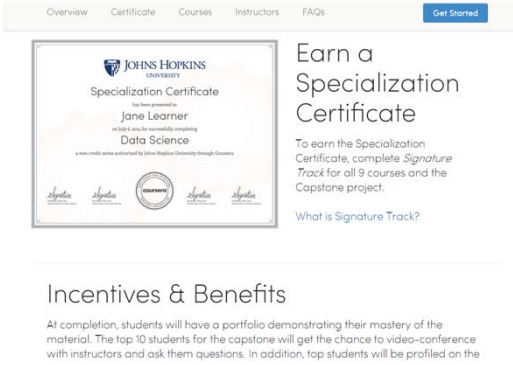
評価実施者のばらつき→参加者のbackgroundは多様
評価者4名の代表性→統計の小標本問題と同様

参加者の集団が**異質**な大標本→
- 相互評価の人数を増やす
- コースの登録時のアンケート回答情報利用→
- 集団をbackgroundにより層別化して各層ごとに評価者を決める等

9 + Project



Certification



Wrap up email

最終評価が出るまで約2週間かかる

Coursera 内の Johns Hopkins Univ. 提供ｺｰｽに注目しておいてほしい

Specialization を修得すれば，他のﾃﾞｰﾀｻｲｴﾝｽｺｰｽや統計学，ﾃﾞｰﾀ解析ｺｰｽに有効

Toolbox ｺｰｽは，renewal してまたすぐに始まるので，討論中の受講者等は再び参加して続けてほしい

テキストの紹介 ebook

The Elements of Data Analytic Style

that covers a **ton of material** from all 9 courses in the Data Science Specialization ($10)

1. 10 things statistics taught us about big data analysis
2. The Leek Group Guide to R packages
3. How to share data with a statistician

JMOOC

ga031: 社会人のためのデータサイエンス入門
　　3月17日開講

ga009: デジタルアーカイブのつくり方
　　次回未定

統計学Ⅰ：データ分析の基礎
　　次回未定

まとめと今後

The Data Scientist's Tool Box (Coursera) の単位認定

Quiz + course project の中身と評価方法

Peer assessment の実際

Peer assessment の課題

※ MOOC主催側→ﾃﾞｰﾀを収集→分析→改善

他のﾃﾞｰﾀｻｲｴﾝｽ関係，統計学関係の認定過程

他のMOOC

評価方法・考え方の比較

謝辞

統計数理研究所・共同研究利用・重点型研究（26-共研-4305，26-共研-4306)および共同研究集会（26-共研-5018）の助成