

データで見つけた ~日本人も知らない日本~ Japan Tourism Data Analytics

外国人に人気の観光地のデータを分析 人気の県のランキング上位について気象データ、アニメ聖地数、食べ物、施設、温泉などを調査し 各県ごとの特徴を発見、日本人も知らない各県の魅力を発見しました!

> 荒井(リコー)、米澤(日立)、入村(インフォシス) 大原、宮城



イントロ

【経緯】

当初は気象条件に合わせた旅行先やレジャーの提案を考えていた。

データ量とデータの正確性の両面から厳しいと分かった。

そこで、観光庁にデータが揃っている訪日外国 人に絞ることにした。



イントロ

【目標】

訪日外国人が快適に旅行するために、オススメ の観光地を様々な側面から発掘する。

- 気象条件
- 食事
- ▶施設
- •文化
- •wifi設備
- ・・・など



データ分析の方針

「外国人延べ宿泊者数」=「人気が高い/お奨めの観光地」と定義して、次の観点でどのあたりが人気の原因かをデータで分析

- 対象地域
- →外国人観光客の宿泊延人数ランキング上位の県
- 各県ごとの特徴
- →外国人観光客の満足度が高い要素との相関を調査
- 季節ごとのおすすめ
- →不快指数との相関で、気候との関係を調査



県別の特徴の分析

外国人観光客の満足度の高い要素について施設数を調査

- 日本食を食べる(寿司、ラーメン、そば)→店舗数
- 自然・景勝地観光、スポーツ等→アウトドア施設数
- 美術館・博物館、舞台鑑賞等→インドア施設数
- 映画・アニメ縁の地を訪問→アニメ聖地数
- 温泉入浴→温泉施設数その他、Wi-Fi、鉄道の駅数などを追加

<分析方法>

- ① 施設数と外国人延べ宿泊者数の相関
- ② 人口1万人あたりの施設数と外国人延べ宿泊者数の相関



データの出典

- 外国人観光客数ランキング:
- 外国人観光客の満足度の高い要素: 「訪日外国人消費動向調査結果及び分析平成29年7-9月 期報告書」(国土交通省観光庁)「4.(4)日本滞在中の行動」
- 飲食店舗数の対象(寿司、ラーメン、そば):
 同上「6.もっとも満足した飲食」
- 飲食施設数
- アウトドア施設数
- インドア施設数
- アニメ聖地数
- 温泉施設数
- フリー無線LANアクセスポイント数







— の前に

突然ですが ここでクイズです



クイズ(1)

アニメ聖地数が一番多い都道府県は東京都です。

では、2番目にアニメ聖地数が多い県は?

①千葉 ②神奈川 ③京都



答え①

1位 東京 4,305 2位 神奈川 1,045 3位 京都 880

4位 北海道

5位 千葉

623

428





クイズ(2)

人口当たりの飲食店数が多い県は?

- ①酪農と農業の一大産地 北海道
- ②甲府ワインが絶品 山梨
- ③なんといっても食い倒れ 大阪



答え②

	寿	司	ラー	-メン	そば					
	都道府県 名	人口1万 人当たり の店舗数	都道府県 名	人口1万 人当たり の店舗数	都道府県 名	人口1万 人当たり の店舗数				
1位	東京	4.1	北海道	6.2	山梨	9.0				
2位	山梨	4.0	山梨	5.4	沖縄	7.0				
3位	大阪	3.5	東京	4.8	東京	4.9				



クイズ(3)

インドア施設数が一番多い県は東京ですが、2番目に多い県は?

①北海道 ②神奈川 ③大阪



答え①

1位 東京 3,281 2位 北海道 1,859 - WEEN -

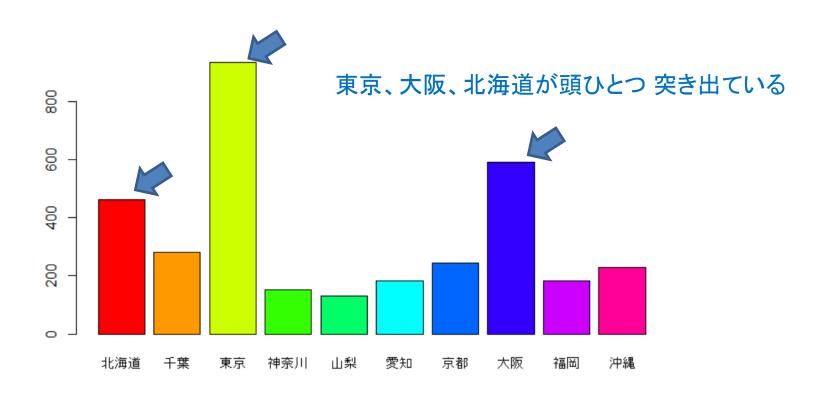






分析結果 - 概要 -

訪日外国人宿泊延人数(2016年)・・・人気度の指標



R Commander でヒストグラム表示

barplot(dataset[,c("宿泊延人数")], names.arg=dataset[,c("都道府県")], col = rainbow(10))



仮説

外国人観光客は、日本らしさを求めて 訪問地を決めているのではないか?

- •食事
- •施設
- •文化
- •WiFi などインフラ



正の相関:そば、寿司、インドア施設、ラーメン、アニメ・・・

C01 寿司	C02ラーメ	C03そば
店数	ン店数	店数
0.87	0.82	0.89











	訪日外国 人宿泊延 人数	C01 寿司 店数	C02ラーメ ン店数	C03そば 店数	CO4アウト ドア施設 数	C05インド ア施設数	C06アニメ 聖地数	C07世界 遺産数	C08神社 数	C09寺社 数	C10教会 数	C11 都市 公園数	C12都市 公園面積	C13博物 館数	C14博物館外国人向け案内	C15温泉 数			C18ホテ ル数	C19旅館 数	C20国際 線空港数	C21 鉄道 駅数	C22WiFi
訪日外国人宿泊延人数	1.00	0.87	0.82	0.89	-0.01	0.87	0.79	0.33	-0.27	0.23	0.75	0.46	0.28	0.47	0.68	0.12	-0.19	0.15	0.66	-0.28	0.61	0.78	0.54
のの対しの	0.07	1.00	0.50	0.55	0.10	0.01	0.02		0.01	0.44	0.00	0.50	0.17	0.01		0.00	0.11	0.00	0.07	0.51	0.44	0.04	0.00
002ラーメン店数	0.82	0.95	1.00	0.94	0.12	0.97	0.85	0.26	0.04	0.36	0.73	0.63	0.36	0.70		0.19	0.10	0.33	0.84		0.57	0.94	0.47
C03そば店数	0.89	0.99	0.94	1.00	-0.16	0.91	0.84	0.27	0.00	0.38	0.81	0.49	0.14	0.50	0.73	-0.06	-0.18	0.05	0.65	-0.38	0.48	0.92	0.66
C04アウトドア施設数	-0.01	-0.15	0.12	-0.16	1.00	0.21	0.00	0.20	-0.24	-0.31	-0.19	0.20	0.73	0.62		0.80	0.74	0.89	0.48	0.64	0.54	-0.02	-0.53
005インドア施設数	0.87	0.91	0.97	0.91	0.21	1.00	0.90	0.31	-0.12	0.27	0.68	0.54	0.40	0.74		0.23	0.11	0.35	0.84	-0.05	0.64	0.88	0.37
C06アニメ聖地数	0.79	0.82	0.85	0.84	0.00	0.90	1.00	0.45	-0.12	0.12	0.47	0.24	0.01	0.46	0.84	-0.09	-0.06	0.05	0.63	-0.14	0.34	0.70	0.22
C07世界遺産数	0.33	0.15	0.26	0.27	0.20	0.31	0.45	1.00	-0.01	-0.30		-0.30	-0.04	0.23	0.27	0.00	-0.21	0.07	0.21	-0.14	0.40	0.10	-0.08
C08神社数	-0.27	0.01	0.04	0.00	-0.24	-0.12	-0.12	-0.01	1.00	0.62	0.05	0.20	-0.12	0.02	-0.23	-0.03	-0.22	-0.10	0.03	-0.07	-0.24	0.17	0.16
C09寺社数	0.23	0.44	0.36	0.38	-0.31	0.27	0.12	-0.30	0.62	1.00	0.63	0.48	0.19	0.34	0.18	0.09	-0.14	0.01	0.20		0.06	0.62	0.63
C10教会数	0.75	0.83	0.73	0.81	-0.19	0.68	0.47	0.03	0.05	0.63		0.62	0.37	0.51	0.49	0.15	-0.08	0.17	0.50	-0.20	0.53	0.89	0.90
C11 都市公園数	0.46	0.56	0.63	0.49	0.20	0.54	0.24	-0.30	0.20	0.48		1.00	0.65	0.55	0.32	0.61	0.35	0.56	0.79	0.06	0.43	0.71	0.41
C12都市公園面積	0.28	0.17	0.36	0.14	0.73	0.40	0.01	-0.04	-0.12	0.19		0.65	1.00			0.93	0.59	0.87	0.68	0.47	0.79	0.40	0.06
C13博物館数	0.47	0.51	0.70	0.50	0.62	0.74	0.46	0.23	0.02	0.34	0.51	0.55	0.82	1.00		0.67	0.49	0.74	0.79	0.41	0.81	0.68	0.16
C14博物館外国人向け案	0.68	0.73	0.83	0.73	0.32	0.91	0.84	0.27	-0.23	0.18		0.32	0.40	0.79		0.20	0.28	0.36	0.70		0.58	0.72	0.16
C15温泉数	0.12	-0.03	0.19	-0.06	0.80	0.23	-0.09	0.00	-0.03	0.09		0.61	0.93	0.67		1.00		0.92	0.62	0.59	0.62	0.21	-0.17
C16温泉宿泊施設数	-0.19	-0.11	0.10	-0.18	0.74	0.11	-0.06	-0.21	-0.22	-0.14	-0.08	0.35	0.59	0.49	0.28	0.65	1.00	0.81	0.35	0.78	0.18	0.04	-0.38
C17温泉公衆浴場数	0.15	0.09	0.33	0.05	0.89	0.35	0.05	0.07	-0.10	0.01	0.17	0.56	0.87	0.74	0.36	0.92	0.81	1.00	0.65	0.64	0.59	0.29	-0.19
C18ホテル数	0.66	0.67	0.84	0.65	0.48	0.84	0.63	0.21	0.03	0.20	0.50	0.79	0.68	0.79	0.70	0.62	0.35	0.65	1.00	0.10	0.69	0.75	0.14
C19旅館数	-0.28	-0.31	-0.14	-0.38	0.64	-0.05	-0.14	-0.14	-0.07	0.08		0.06	0.47	0.41	0.19	0.59	0.78	0.64	0.10		0.04	-0.12	-0.47
C20国際線空港数	0.61	0.44	0.57	0.48	0.54	0.64	0.34	0.40	-0.24	0.06	0.53	0.43	0.79	0.81		0.62	0.18	0.59	0.69	0.04	1.00	0.54	0.27
C21 鉄道駅数	0.78	0.94	0.94	0.92	-0.02	0.88	0.70	0.10	0.17	0.62	0.89	0.71	0.40	0.68	0.72	0.21	0.04	0.29	0.75	-0.12	0.54	1.00	0.69
C22WiFi	0.54	0.66	0.47	0.66	-0.53	0.37	0.22	-0.08	0.16	0.63	0.90	0.41	0.06	0.16	0.16	-0.17	-0.38	-0.19	0.14	-0.47	0.27	0.69	1.00



仮説

気候の良い春や秋は日本中が観光シーズン! 外国人の観光する季節も 「不快指数」と関係あるのではないか?



ところが一

宿泊延数と気候の間に明確な相関関係は見えてこない 快適な気候と観光先は関係がない?

	観光目的_ 宿泊延数	不快指 数	平均気温	平均湿 度	降水日 数	月間降 水量	暑い日	寒い日	月間降雪量	風の日	雪日数	雷日数	霧日数	平均風速	日照時間
観光目的_宿泊延数	1.00	-0.03	-0.03	0.03	0.09	0.01	-0.01	0.02	0.10	-0.15	0.09	-0.02	-0.17	-0.12	-0.07
不快指数	-0.03	1.00	1.00	0.01	0.21	0.49	0.72	0.03	0.48	-0.07	0.05	0.55	0.30	0.10	0.18
平均気温	-0.03	1.00	1.00	0.60	0.21	0.48	0.71	-0.64	-0.48	-0.07	-0.65	0.54	-0.31	0.15	0.18
平均湿度	0.03	0.61	0.60	1.00	0.60	0.59	0.54	-0.23	-0.04	-0.05	-0.15	0.43	-0.22	0.17	-0.39
降水日数	0.09	0.21	0.21	0.60	1.00	0.58	0.27	0.12	0.30	0.10	0.25	0.40	-0.20	0.11	-0.52
月間降水量	0.01	0.49	0.48	0.59	0.58	1.00	0.45	-0.19	-0.06	-0.03	-0.17	0.50	-0.13	0.11	-0.23
暑い日	-0.01	0.72	0.71	0.54	0.27	0.45	1.00	-0.21	-0.14	0.00	-0.22	0.65	-0.14	0.17	0.14
寒い日	0.02	-0.63	-0.64	-0.23	0.12	-0.19	-0.21	1.00	0.79	0.17	0.85	-0.18	0.38	-0.06	-0.15
月間降雪量	0.10	-0.48	-0.48	-0.04	0.30	-0.06	-0.14	0.79	1.00	0.12	0.85	-0.09	0.13	0.00	-0.27
風の日	-0.15	-0.07	-0.07	-0.05	0.10	-0.03	0.00	0.17	0.12	1.00	0.16	-0.01	0.08	0.74	0.00
雪日数	0.09	-0.65	-0.65	-0.15	0.25	-0.17	-0.22	0.85	0.85	0.16	1.00	-0.17	0.11	0.01	-0.31
雷日数	-0.02	0.55	0.54	0.43	0.40	0.50	0.65	-0.18	-0.09	-0.01	-0.17	1.00	-0.14	0.08	0.05
霧日数	-0.17	-0.30	-0.31	-0.22	-0.20	-0.13	-0.14	0.38	0.13	0.08	0.11	-0.14	1.00	-0.06	0.07
平均風速	-0.12	0.16	0.15	0.17	0.11	0.11	0.17	-0.06	0.00	0.74	0.01	0.08	-0.06	1.00	-0.03
日照時間	-0.07	0.18	0.18	-0.39	-0.52	-0.23	0.14	-0.15	-0.27	0.00	-0.31	0.05	0.07	-0.03	1.00



閑話休題

不快指数とは?



閑話休題

不快指数とは?

実は気象庁が決めたものではないらしい。





不快指数

不快指数とは夏の蒸し暑さを数量的に表した指数。 1957年、アメリカ合衆国で考案された。

ヒトは発汗機能が非常によく発達した動物であり、基本的に暑熱に対する耐性が強い。汗による体温調整は、主に汗が皮膚表面で蒸発する際に気化熱を奪うことによってなされる。しかし湿度が高い状態においては、汗の蒸発が阻害されるため「蒸し暑い」と感じる。

Wikipedia より

https://ja.wikipedia.org/wiki/%E4%B8%8D%E5%BF%AB%E6%8C%87%E6%95%B0



不快指数

不快指数とは夏の蒸し暑さを数量的に表した指数。 1957年、アメリカ合衆国で考案された。

「寒くて乾燥」は本来は関係がない?

ヒトは発汗機能が非常によく発達した動物であり、基本的に暑熱に対する耐性が強い。汗による体温調整は、主に<u>汗が皮膚表面で蒸発する際に気化熱</u>

を奪うことによってなされる。しかし湿原

が阻害されるため「蒸し暑い」と感じる

Wikipedia より

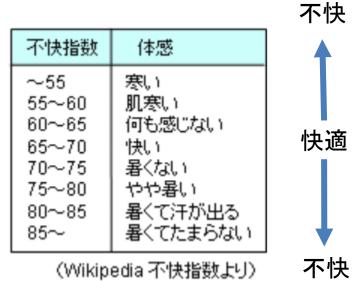
https://ja.wikipedia.org/wiki/%E4%B8%8D%





不快指数





http://keisan.casio.jp/has10/SpecExec.cgi?id=system/2006/1202883065



不快指数の問題 数値が中間ぐらいだと快適 数値が大きくても、小さくても不快 相関関係で扱いづらい

「快適指数」を勝手に作ってみた大きな数値なら快適 (1.0 が最高値)小さな数値だと不快 (0∼ -0.5 ぐらいが最小値)



快適指数:これでも相関があるようには見えない

正の相関:気象以外では、

「お祭り」は ゆる~い 相関がありそう

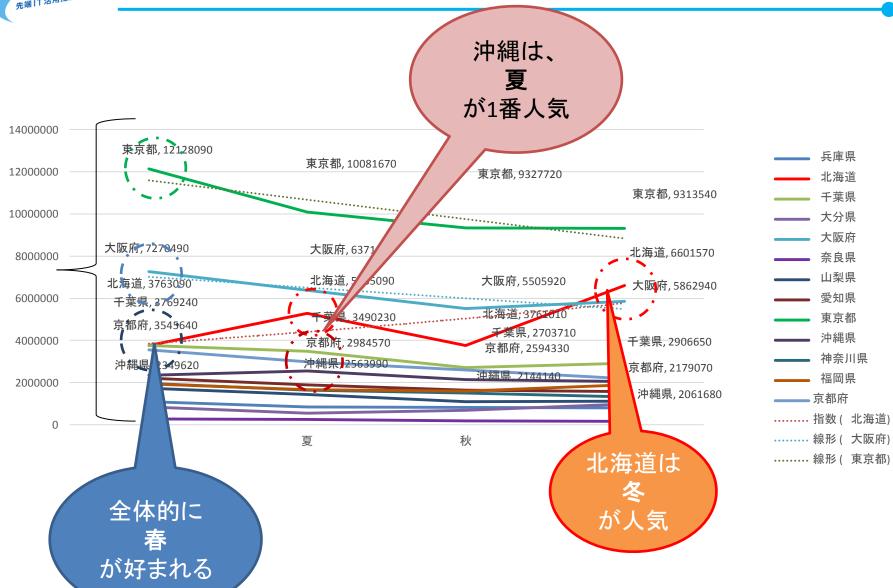
																		1
	観光目的_ 宿泊延数	快適	不快指 数	平均気温	平均湿度	降水日 数	月間降 水量	暑い日	寒い日	月間降雪量	風の日	雪日数	雷日数	霧日数	平均風速	日照時間	お祭り数	果物狩 り
観光目的_宿泊3	1.00	-0.03	-0.03	-0.03	0.03	0.09	0.01	-0.01	0.02	0.10	-0.15	0.09	-0.02	-0.17	-0.12	-0.07	0.45	-0.21
快適	-0.03	1.00	0.74	0.76	0.30	0.01	0.23	0.09	-0.73	-0.57	-0.11	-0.74	0.15	-0.30	0.05	0.16	0.38	0.26
不快指数	-0.03	U.74	1.00	1.00	0.61	0.21	0.49	0.72	-0.63	-0.48	-0.07	-0.65	0.55	-0.30	0.16	0.18	0.30	0.20
平均気温	-0.03	0.76	1.00	1.00	0.60	0.21	0.48	0.71	-0.64	-0.48	-0.07	-0.65	0.54	-0.31	0.15	0.18	0.30	0.21
平均湿度	0.03	0.30	0.61	0.60	1.00	0.60	0.59	0.54	-0.23	-0.04	-0.05	-0.15	0.43	-0.22	0.17	-0.39	0.19	-0.05
降水日数	0.09	0.01	0.21	0.21	0.60	1.00	0.58	0.27	0.12	0.30	0.10	0.25	0.40	-0.20	0.11	-0.52	0.03	-0.05
月間降水量	0.01	0.23	0.49	0.48	0.59	0.58	1.00	0.45	-0.19	-0.06	-0.03	-0.17	0.50	-0.13	0.11	-0.23	0.10	-0.05
暑い日	-0.01	0.09	0.72	0.71	0.54	0.27	0.45	1.00	-0.21	-0.14	0.00	-0.22	0.65	-0.14	0.17	0.14	0.10	0.06
寒い日	0.02	-0.73	-0.63	-0.64	-0.23	0.12	-0.19	-0.21	1.00	0.79	0.17	0.85	-0.18	0.38	-0.06	-0.15	-0.25	-0.13
月間降雪量	0.10	-0.57	-0.48	-0.48	-0.04	0.30	-0.06	-0.14	0.79	1.00	0.12	0.85	-0.09	0.13	0.00	-0.27	-0.17	-0.13
風の日	-0.15	-0.11	-0.07	-0.07	-0.05	0.10	-0.03	0.00	0.17	0.12	1.00	0.16	-0.01	0.08	0.74	0.00	-0.21	-0.15
雪日数	0.09	-0.74	-0.65	-0.65	-0.15	0.25	-0.17	-0.22	0.85	0.85	0.16	1.00	-0.17	0.11	0.01	-0.31	-0.22	-0.18
雷日数	-0.02	0.15	0.55	0.54	0.43	0.40	0.50	0.65	-0.18	-0.09	-0.01	-0.17	1.00	-0.14	0.08	0.05	0.11	0.15
霧日数	-0.17	-0.30	-0.30	-0.31	-0.22	-0.20	-0.13	-0.14	0.38	0.13	0.08	0.11	-0.14	1.00	-0.06	0.07	-0.18	-0.02
平均風速	-0.12	0.05	0.16	0.15	0.17	0.11	0.11	0.17	-0.06	0.00	0.74	0.01	0.08	-0.06	1.00	-0.03	-0.11	-0.21
日照時間	-0.07	0.16	0.18	0.18	-0.39	-0.52	-0.23	0.14	-0.15	-0.27	0.00	-0.31	0.05	0.07	-0.03	1.00	0.19	0.21
お祭り数	0.45	0.36	0.30	0.30	0.19	0.03	0.10	0.10	-0.25	-0.17	-0.21	-0.22	0.11	-0.18	-0.11	0.19	1.00	0.12
果物狩り	-0.21	0.26	0.20	0.21	-0.05	-0.05	-0.05	0.06	-0.13	-0.13	-0.15	-0.18	0.15	-0.02	-0.21	0.21	0.12	1.00



外国人観光客は気候なんて気にしないのか?



AHC 季節ごとに人気の観光地





HC 分析結果から言えること

海外旅行者向けおすすめの県

- ・山梨:食べ物が充実(人口あたりの飲食店)
- WiFi環境は西日本の方が整備されている
- ・ アニメ聖地: やっぱり東京、次点は京都
- インドア:北海道、沖縄
- アウトドア:山梨



苦労話

気象データのダウンロード方法





苦労した点

一人で作業に着手しようとすたところ、どのデータをどのようにダウンロードするべきか迷いました。 チームで取り組み、気象庁の方の説明の元、データ項目をを選択し、まとめてダウンロードしました。 → 勉強会で気象庁の方のアドバイスとチームでの取り組めたのため、大変効率が良かったです。

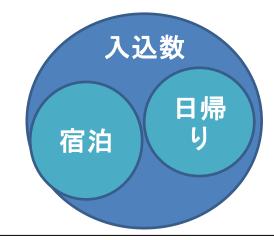


苦労話

宿泊延べ人数・入込数の違い

- → 宿泊延べ人数は(人数x宿泊した日数)で算出されます。
- → 入込数には宿泊と日帰りの観光客の両方が含まれます!
- この2つは市場調査する上で規模を把握するのに大切な2つの用語です。





苦労した点

各都道府県のウエブサイトからデータをダウンロードを試みましたが、データが揃っていない地域が一部ありました。都庁、県庁にも連絡したものの、データが揃いませんでした。

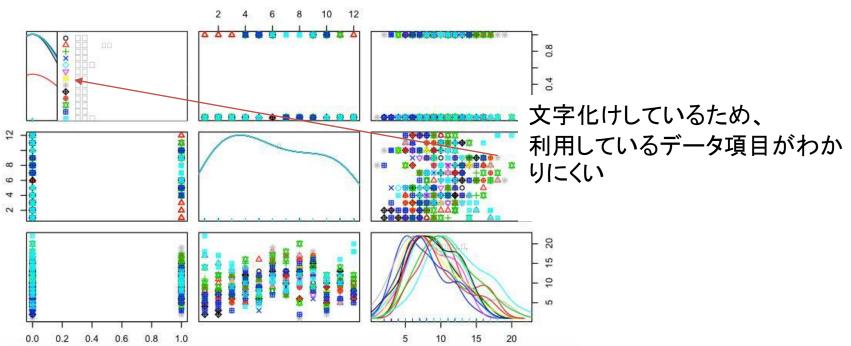


苦労話

文字化け

文字化けが発生したのでデータ分析をExcelで実施したり、異なるPCで作業するなどしました。

原因: Excel が出力した文字コード (Shift JIS)をR に読み込ませるときに UTF-8 に変換してしまうことが判明しています。





その他技術的な点

• CSVよりExcel

データがExcel形式でないと、Rがデータのファイルを認識しなかったので、Excelを使いました。

• MACとWindowsではOffice製品が違う

チーム内のデータやファイルのやり取りで、トラブル が多発



R関数

- 1. Rコマンダー
- 2. 散布図 Scatterplot
- 3. 相関行列 Correlation Matrix

インストールトラブル:

R-3.4.1をインストールしていないと、Rコマンダーをダウンロードできないことが判明、数名がRのバージョンが古いため、Rのバージョンをアップデートしました。



開発環境



・ <u>苦労した点</u>

一部メンバーは、安い作業用のPCを**秋葉原**まで買いに行きました。

→オープンソース界隈の人はとてもお世話好き、秋葉原のお店を案内していただきました。

→オープンソースは便利で、高度なことができます。

トラブルに巻き込まれても解決するのにそれなりの技術力と知識が必要です。



最新情報は こちらをご参照ください

http://aitc.jp

https://www.facebook.com/aitc.jp



ハルミン AITC非公式イメージキャラクター