

Team Ryukyuが3年連続優秀賞から 最優秀賞を獲得！！

兼城宇基¹，高良宙矢²，玉那覇詩乃²，岩永良太³，
稲嶺穂乃花³，山田健太^{1,2}

¹ 琉球大学 地域共創研究科，² 琉球大学 国際地域創造学部

³ 株式会社ビーンズラボ

1. データ解析コンペティションとは（開催体制）

【共催団体】

経営科学系研究部会連合協議会

(Joint Association Study Group of Management Science)

- 日本オペレーションズ・リサーチ ビッグデータを用いたマーケティング分析研究部会
- 日本 マーケティング・サイエンス学会 ID 付 POS データ活用研究部会
- 日本マーケティング・サイエンス学会 消費者・市場反応の科学的研究部会
- 日本マーケティング・サイエンス学会 消費者行動の学際的研究部会
- 日本マーケティング・サイエンス学会 市場予測のための消費者行動分析研究部会
- 日本計算機統計学会 データ解析スタディーグループ
- 日本経営工学会 経営情報部門
- 株式会社NTTデータグループ 技術革新統括本部 技術開発本部
- 株式会社産業科学研究開発センター

<共 催 : >

- 日本ソーシャルデータサイエンス学会

1. データ解析コンペティション（趣旨）

【趣旨】

- コンペティションを通じて、データ解析及びモデル設計を行うことで、新たな知見を誘発すること
- お互いの発表を通じてデータ分析、データ活用に関する議論を活発にすること
- これらを通じて、学術的及び実務的な研究の発展に寄与すること

【基本コンセプト】

- 産学両方から幅広く参加チームを募集します
- 共通のデータを提供し、それぞれの視点で分析をします
- データ分析の新規性、ビジネス有効性を競います
- すべてのチームが成果発表をし、切磋琢磨します

1. データ解析コンペティション（過去のテーマとデータ）

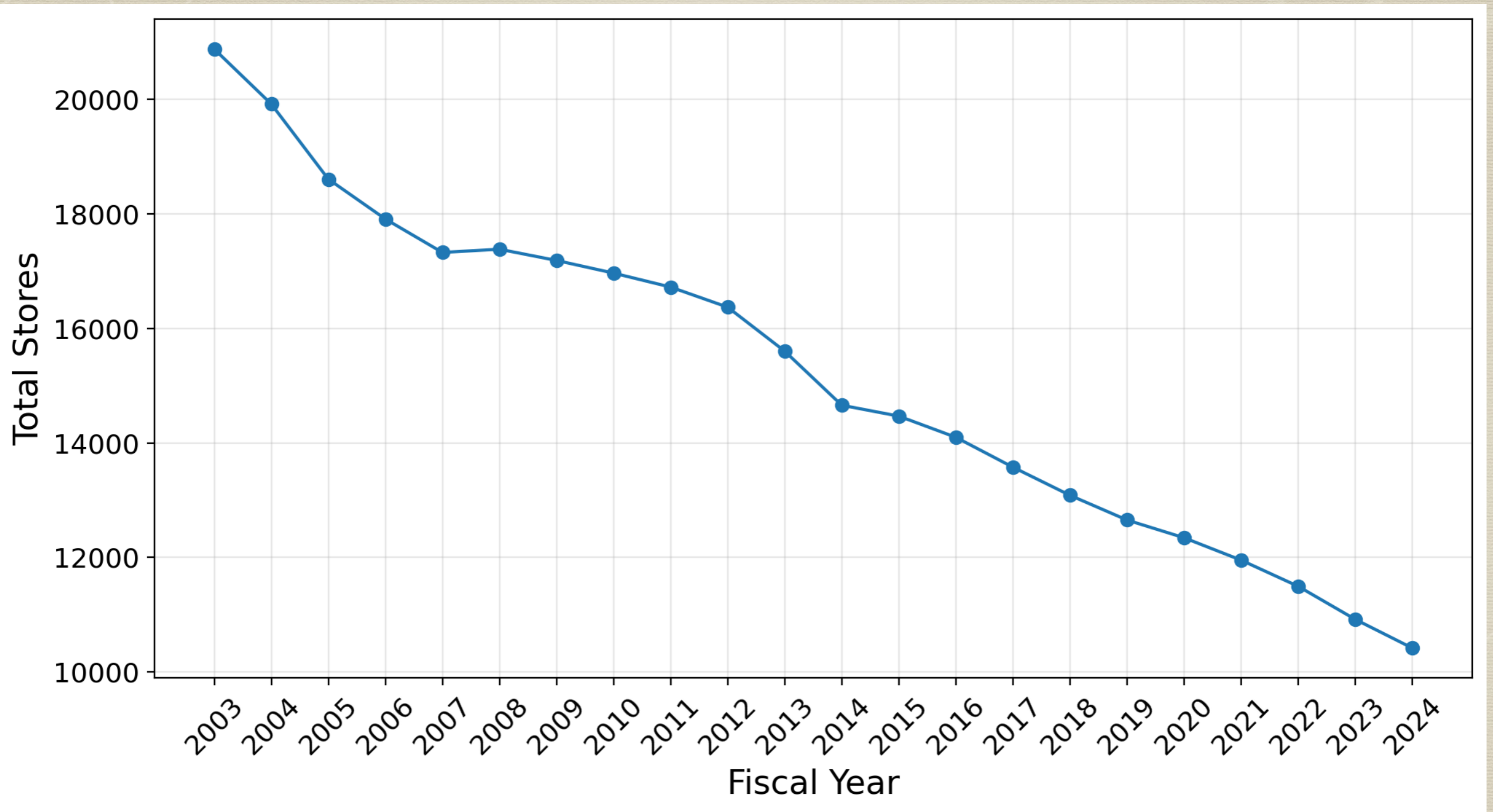
年度	データ	年度	データ	年度	データ
H6年度	<u>食品購買行動</u> ● ストア・スキャン・データ	H17年度	<u>アミューズメントPOS</u> ● CD販売店ID付POS	H28年度	<u>ファッションEC</u> ● ECのID-POS, 生活意識調査
H7年度	<u>食品・日雑購買行動</u> ● ストア・スキャン・データ	H18年度	<u>WEBマーケティング</u> ● WEBアクセスログ	H29年度	<u>サービス産業ID-POS</u> ● ヘアサロンのID-POSデータ
H8年度	<u>日用雑貨品購買行動</u> ● ホーム・スキャン・データ	H19年度	<u>オークション・データ分析</u> ● B2B 自動車オークション・データ	H30年度	<u>生活者のメディア接触分析</u> ● メディ視聴・接触データ
H9年度	<u>観光行動</u> ● 旅行履歴, 意識アンケート	H20年度	<u>消費場面分析</u> ● 食卓メニューデータ	R元年度	<u>交通データの分析</u> タクシー・プローブ・データ
H10年度	<u>食卓マーケティング</u> ● メニュー・データ	H21年度	<u>百貨店分析</u> ● 百貨店ID-POSデータ	R2年度	<u>大規模生活者アンケートの分析</u> 多数の生活者アンケート
H11年度	<u>金融マーケティング</u> ● 行動, 意識アンケート	H22年度	<u>日用品ID付POS</u> ● ドラッグストア・ID付POS	R3年度	<u>Q&Aサイトデータ</u> 妊娠・出産・子育てママ対象Q&Aサイトデータ
H12年度	<u>金融マーケティング</u> ● 銀行取引サマリ ● 行動, 意識アンケート	H23年度	<u>ウェブ・マーケティング</u> ● ウェブアクセス+購買履歴	R4年度	<u>ECサイト販売履歴データ</u> 大手ECサイト販売履歴データ
H13年度	<u>流通CRM</u> ● ポイントカード・データ	H24年度	<u>サービス・マーケティング</u> ● クーポン共同購入サイト ● 不動産情報サービスサイト	R5年度	<u>飲料POSデータ</u> 10年分POSデータ
H14年度	<u>流通CRM</u> ● ポイントカード・データ	H25年度	<u>消費者行動分析</u> ● ECアクセス・購買ログ・データ ● ホーム・スキャン・データ	R6年度	<u>食品レシートデータ</u> レシート読み取りアプリデータ
H15年度	<u>10周年記念</u> ● 電力消費, ハウスカード, スーパー・ドラッグストアPOS	H26年度	<u>新たな顧客接点</u> ● 小売業FSPデータ+POS ● ID-POS, EC購買履歴, アプリ利用		
H16年度	<u>2種類のデータを提供</u> ● クレジットカード利用履歴, 加工食品POS	H27年度	<u>データの新たな展望</u> ● 複数チェーンID-POS ● 行政窓口受付データ		

データ概要 (令和7年度コンペ)

- 35店舗の書店での書籍・雑誌販売データ
- 「店舗別」「日別」「書籍」別に集計データ：
提供用_日別販売明細.txt
 トランザクション数：3,969,881
 最大販売数：9,458
 再販維持価格
- 店舗別の購買者数データ：
提供用_店舗日別客数.txt
- 書店コード対応表（取り扱い注意！）：
提供用_書店コード.txt
- 期間は2024年1月1日～12月31日

日付	書店コード	ISBN	出版社	書名	著者名	大分類	中分類	小分類	本体価格	POS販売冊数
0	2024-01-01 00:00:00.000	1 978-4-576-23148-8	二見書房	書院番組頭	藤木桂	文庫	日本文学	二見書房	800.0	1
1	2024-01-01 00:00:00.000	1 978-4-04-737647-2	KADOKAWA	珍獣のお医者さん 1	二宮香乃	コミック	少年(小中学生)	KADOKAWA ハルタC	720.0	1
2	2024-01-01 00:00:00.000	1 978-4-575-52651-6	双葉社	こんな日は喫茶トードーで雨宿り。	標野凧	文庫	日本文学	双葉文庫	630.0	1
3	2024-01-01 00:00:00.000	1 978-4-12-207455-2	中央公論新社	病み虫 うぼっぼ同心十手綴り	坂岡真	文庫	日本文学	中公文庫	700.0	1
4	2024-01-01 00:00:00.000	1 978-4-15-031562-7	早川書房	ピュア	小野美由紀	文庫	日本文学	ハヤカワ文庫	1000.0	1
...
3969876	2024-12-31 00:00:00.000	35 978-4-05-750848-1	Gakken	頭のよくなる算数ゲーム	リ：チェント 稲葉直貴	児童	知育絵本	カード	1700.0	1
3969877	2024-12-31 00:00:00.000	35 978-4-09-942527-2	小学館	シナぷしゅ えあわせカード	NaN	児童	知育絵本	カード	1300.0	1
3969878	2024-12-31 00:00:00.000	35 978-4-8133-2630-4	TTJたちばな	ミッキーマウスとロードレーサーズ	NaN	児童	キャラクター	ディズニー	398.0	1
3969879	2024-12-31 00:00:00.000	35 978-4-7747-3876-5	コスミック出版	きらめく海と珊瑚礁のファンタジー	西脇エリ	趣味	イラスト・カット	イラスト・カットその他	1500.0	1
3969880	2024-12-31 00:00:00.000	35 978-4-88574-642-0	東京書店	キラキラぬりえファンシープリンセス	NaN	児童	知育絵本	ぬりえ・お絵かき	850.0	1

書店数の変化



出版科学研究所のデータより作成

- 書店数は20年間で半減(約20000→約10000店)
- データを用いて作業の効率化や集客ができれば書店に貢献できる

3. コンペの進め方(1) : 全体スケジュール

- 発会式 (8月8日)
- チームエントリー (8月8日から25日まで) →まずは**Excelファイル**で
- 参加の決定とデータの貸与 (9月上旬まで) →**署名捺印された申請書・誓約書送付後**
なお, データのダウンロードにあたり代表者が管理する「**googleアカウント(gmailアドレス, google suiteを使っていれば職場のアカウントでも構いません)**」が必要です.
- (分析ツールの案内を9月上旬頃に予定)
- **10月~3月 (次ページのいずれかの研究部会で発表をしてください)**
 - 12/20- **中間発表 : 互いの分析の視点の違いなどについて議論ください**
 - 2/21- **最終発表 : 各研究部会で審査を行います.**
 - 3/19 **成果報告会への出場チームを決定いただきます.**
- **3月上旬~中旬 (成果報告会・審査・表彰)**

4. 審査の視点(成果報告会)

■ 研究内容についての審査基準

- 学術的な新規性
- 結果の信頼性
- 適用可能性

などを総合的に評価

■ 実用的なメッセージ

- マーケティング・データの分析については、アクションにつながる結論が重視されます

■ プレゼンテーション力

- 研究内容を的確に伝え、質疑応答なども含めたコミュニケーション能力も審査基準とします

2/21の発表会のプログラム

10:00 開会

【学生部門 1】 10:05~12:05

1. K's PANDA ささだんご 大竹恒平 (上智大学) [10+3 分]
2. 竹取物語 大竹恒平 (上智大学) [10+3 分]
3. LGX- α 田畑智章 (東海大学) [7+3 分]
4. LGX- β 田畑智章 (東海大学) [7+3 分]
5. LGX- γ 田畑智章 (東海大学) [7+3 分]
6. 久保田ゼミ 久保田貴文 (多摩大学) [7+3 分]
7. AMP Entrees 久保田貴文 (多摩大学) [7+3 分]
8. AMP Plats 久保田貴文 (多摩大学) [7+3 分]
9. AMP Desserts 久保田貴文 (多摩大学) [7+3 分]
10. ADSL-R7 南弘征 (北海道大学) [10+3 分]

【学生部門 2】 12:45~14:55

11. FONTAINEBLEAU 船山貴光 (西武文理大学) [7+3 分]
12. 理大マネジメント 森裕一 (岡山理科大学) [7+3 分]
13. T-wave2025NG 山本義郎 (東海大学) [5+3 分]
14. T-wave2025B4 山本義郎 (東海大学) [10+3 分]
15. 青山経シスアナリシス 齊藤史哲 (青山学院大学) [10+3 分]
16. S Lab 酒折文武 (中央大学) [7+3 分]
17. Team KASUMI2025 藤野友和 (福岡女子大学) [10+3 分]
18. Honey Bunnies 竹内光悦 (実践女子大学) [7+3 分]
19. 紅茶派 竹内光悦 (実践女子大学) [7+3 分]
20. 行動 C ラボ 2025 竹内光悦 (実践女子大学) [7+3 分]
21. UEC データ工房 2025 高橋里司 (電気通信大学) [10+3 分]

【学生部門 3】 15:05~16:30

22. T-wave2025 山本義郎 (東海大学) [7+3 分]
23. ブラックキッズ+ 白井康之 (大東文化大学) [10+3 分]
24. 焼肉食べたい 朝日弓未 (東京理科大学) [7+3 分]
25. QU エコノクエスト 北原知就 (九州大学) [7+3 分]
26. NaN しか勝たん 後藤裕介 (芝浦工業大学) [10+3 分]
27. チームソウハツ 後藤裕介 (芝浦工業大学) [10+3 分]
28. jasp 山本由和 (徳島文理大学) [10+3 分]

【一般】 16:30~17:20

29. ストーニーゲート 坂本真仁 (金沢工業大学) [10+3 分]
30. チーム廃藩置県 宮田龍 (琉球大学) [10+3 分]
31. TeamRyukyu 山田健太 (琉球大学) [10+3 分]
32. みかかかさ 山田俊哉 (NTT テクノクロス株式会社) [7+3 分]

32チームが参加

5チームが最優秀賞

Team Ryukyuの最近の成績

22年度:大手ECサイトの購買履歴データ (優秀賞)

23年度:スーパーの10年分POSデータ (優秀賞)

24年度:レシート読取アプリのデータ (優秀賞)

25年度:書店の売上データ (最優秀賞) ← 今回

昨年度の学長記者懇談会



沖縄タイムス(2025.4.2)

最優秀賞を獲得できた要因

- 修士課程の学生(兼城さん)が加わり分析力が大幅にUP！
- 昨年度も参加し実力をつけた学生(玉那覇さん, 高良さん)が今年も参加
- データサイエンティスト養成履修カリキュラムで身につけた基礎力
- ビーンズラボの方(岩永さん, 稲嶺さん)からのフィードバック
- 研究結果を活かしたダッシュボードを作成し実務への応用を明確に示した



データサイエンスと3種の構成要素

数理学

数学・統計学
数理モデリング

(ビッグ)データ
分析
モデリング

解析的計算
(手計算)

データ
サイエンス

Python, R, ...
繰り返し, 分岐
関数, ...

(ビッグ)データ
収集

物理・化学
生物・経済・金融
マーケティング
医学・スポーツ
地域・文化

プログラミング

ドメイン知識

+問題発見力, AI力, 工夫力, プレゼン力, コミュ力, 行動力など 12

データサイエンティスト養成履修カリキュラム

リテラシーレベル

修了要件：必修2または4単位
★獲得要件：修了要件+選択2単位以上

【必修科目】

2024以前入学生 2025以降入学生

情報科学演習

2024以前入学生

データサイエンス
概論

データサイエンス
初級

【選択科目】

AI入門

応用基礎レベル

修了要件：必修6単位+数学系科目2単位以上
★獲得要件：修了要件+PBL科目2単位以上

【必修科目】

データサイエンス
概論

データサイエンス
初級

統計科学

データサイエンス
のためのPython
プログラミング

【数学系科目】

社会科学のための
統計（基礎）

データサイエンス
のための基礎数学

社会科学のための
統計（応用）

社会科学のための
数学（基礎）

社会科学のための
数学（応用）

【PBL科目】

データサイエンス
実践演習Ⅰ

データサイエンス
実践演習Ⅳ

データサイエンス
実践演習Ⅱ

データサイエンス
実践演習Ⅴ

データサイエンス
実践演習Ⅲ

データサイエンス
実践演習Ⅵ

応用発展レベル

修了要件：必修2単位+数学系科目4単位以上
+PBL科目2単位以上
★獲得要件：修了要件+選択2単位以上

【必修科目】

データサイエンス
中級

機械学習

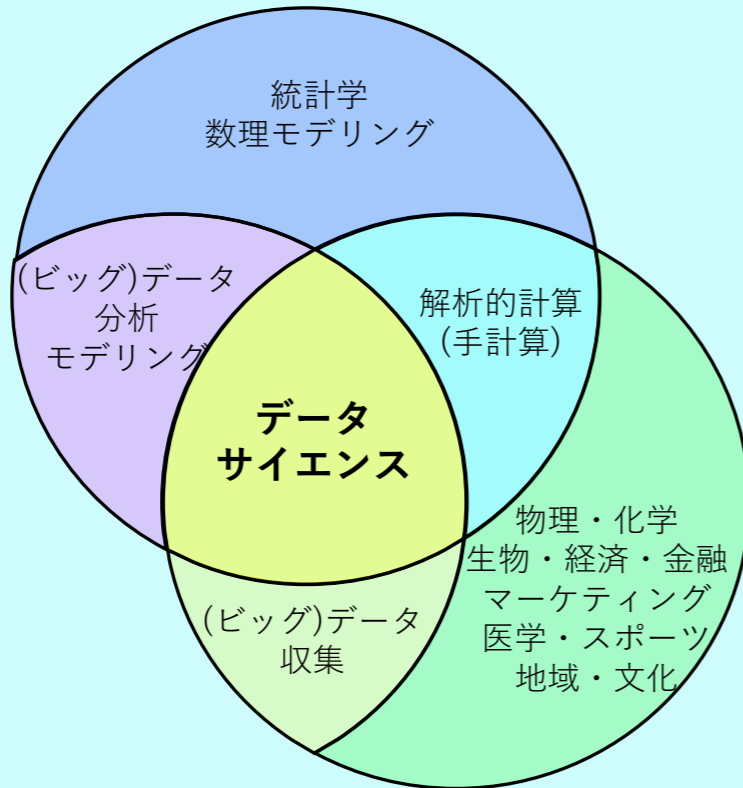
【選択科目】ディープラーニング

データサイエンス
上級

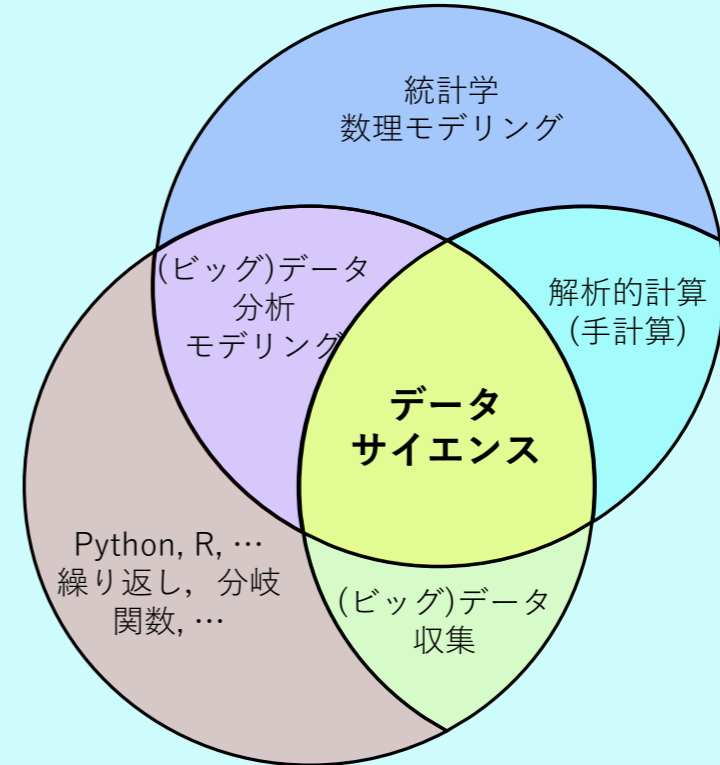


データサイエンス科目のカバー範囲

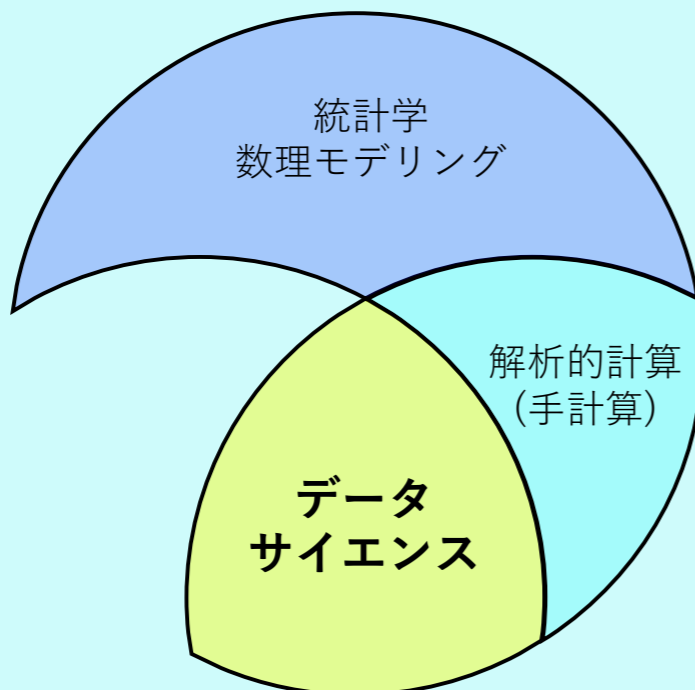
データサイエンス概論



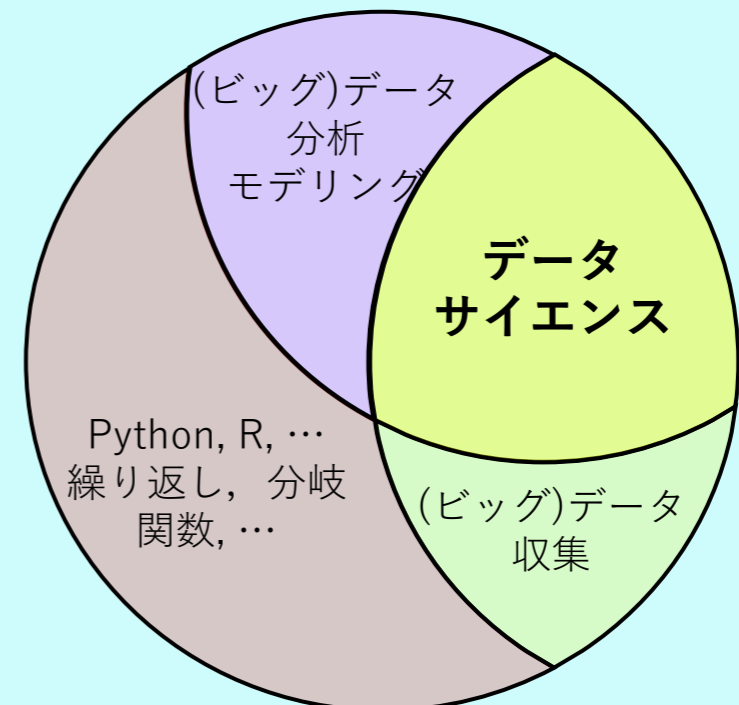
データサイエンス初級(入門)・中級(応用)



データサイエンスのための基礎数学



データサイエンスのためのプログラミング



各科目の難易度

AI・DS人材育成ピラミッド

トップ

100人/年

エキスパート

2000人/年

応用基礎レベル

25万人/年

学部3, 4年

リテラシーレベル

50万人/年

(大学・高専卒業生全員)

学部1, 2年

データサイエンス概論

情報科学演習

AI入門

データサイエンス初級

基礎数学

データサイエンスのための

社会科学のための統計

社会科学のための数学

Pythonプログラミング

データサイエンスのための

データサイエンス実践演習

データサイエンス中級

データサイエンス上級

(ゼミ)

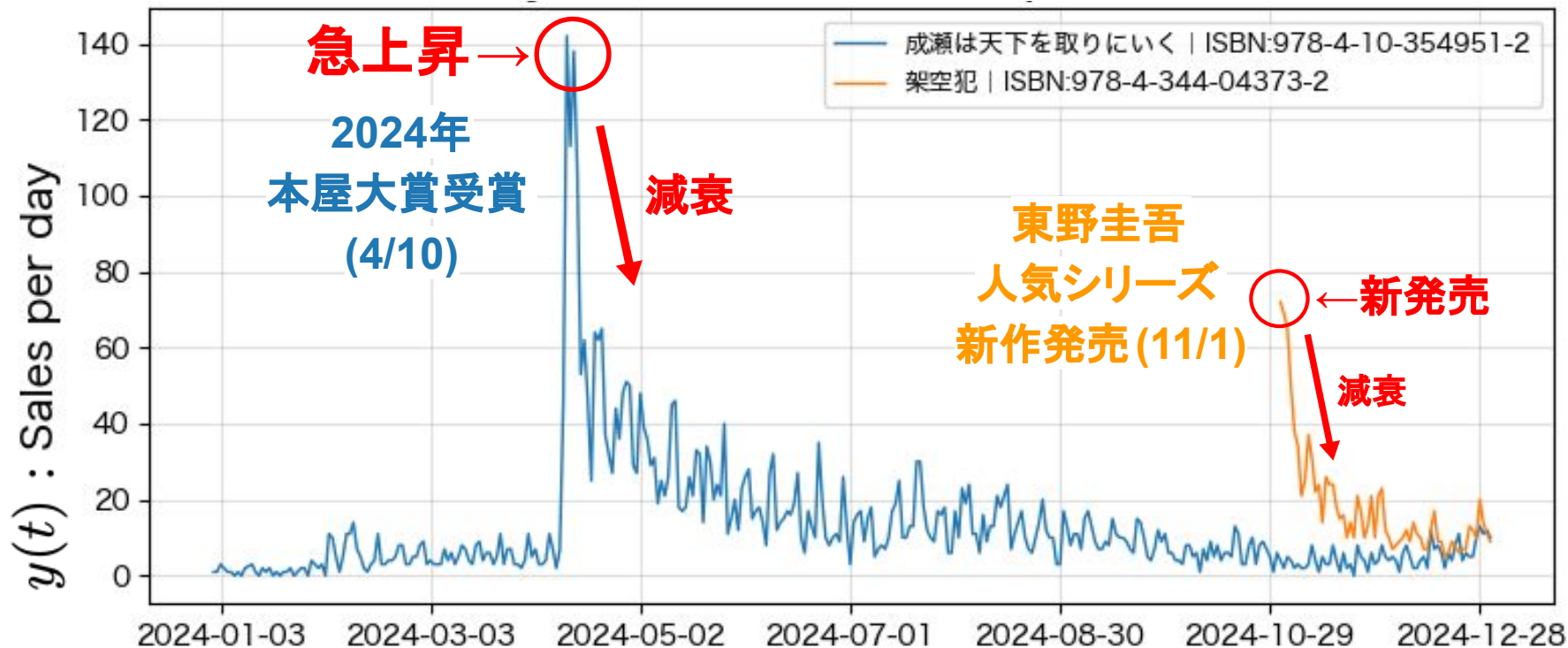
(大学院)

タイトル別販売数の 急上昇検出と減衰のモデル化

兼城宇基¹, 高良宙矢¹, 玉那覇詩乃¹, 岩永良太²,
稲嶺穂乃花², 山田健太¹

¹ 琉球大学国際地域創造学部, ² (株)ビーンズラボ

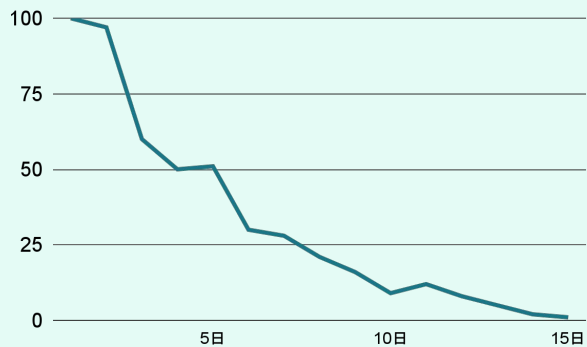
販売冊数の時系列から観測される急上昇と減衰



- ・本屋大賞など外勢的要因によって急激に売り上げが伸びる本がある
- ・急上昇や新発売直後は、基本的に減衰する

ピークのパターン

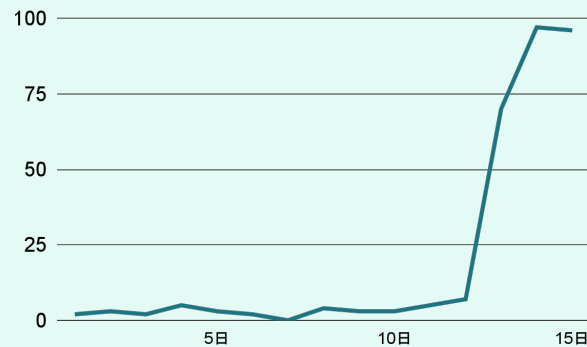
発売日直後



緩やかな上昇



急上昇



新発売日の取得

国立国会図書館のデータベースからAPIを用いて発売年月を取得

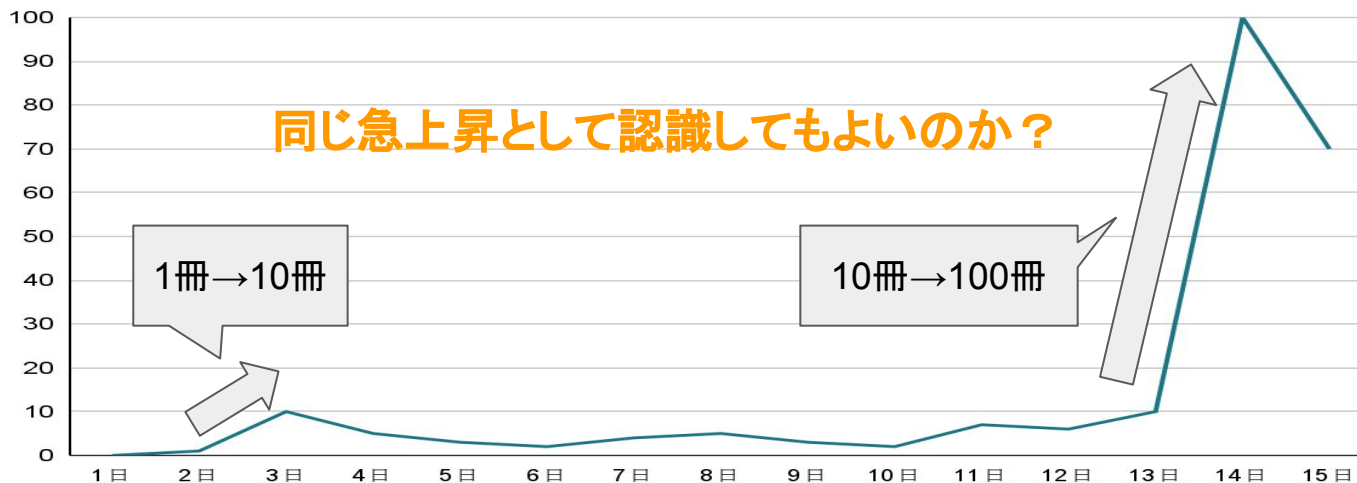
2024年に発売され、コミック・月刊誌・週刊誌で発売時の売上ピークが20冊以上のタイトルを選定

大分類	タイトル数
月刊誌	2012
コミック	1303
週刊誌	1156



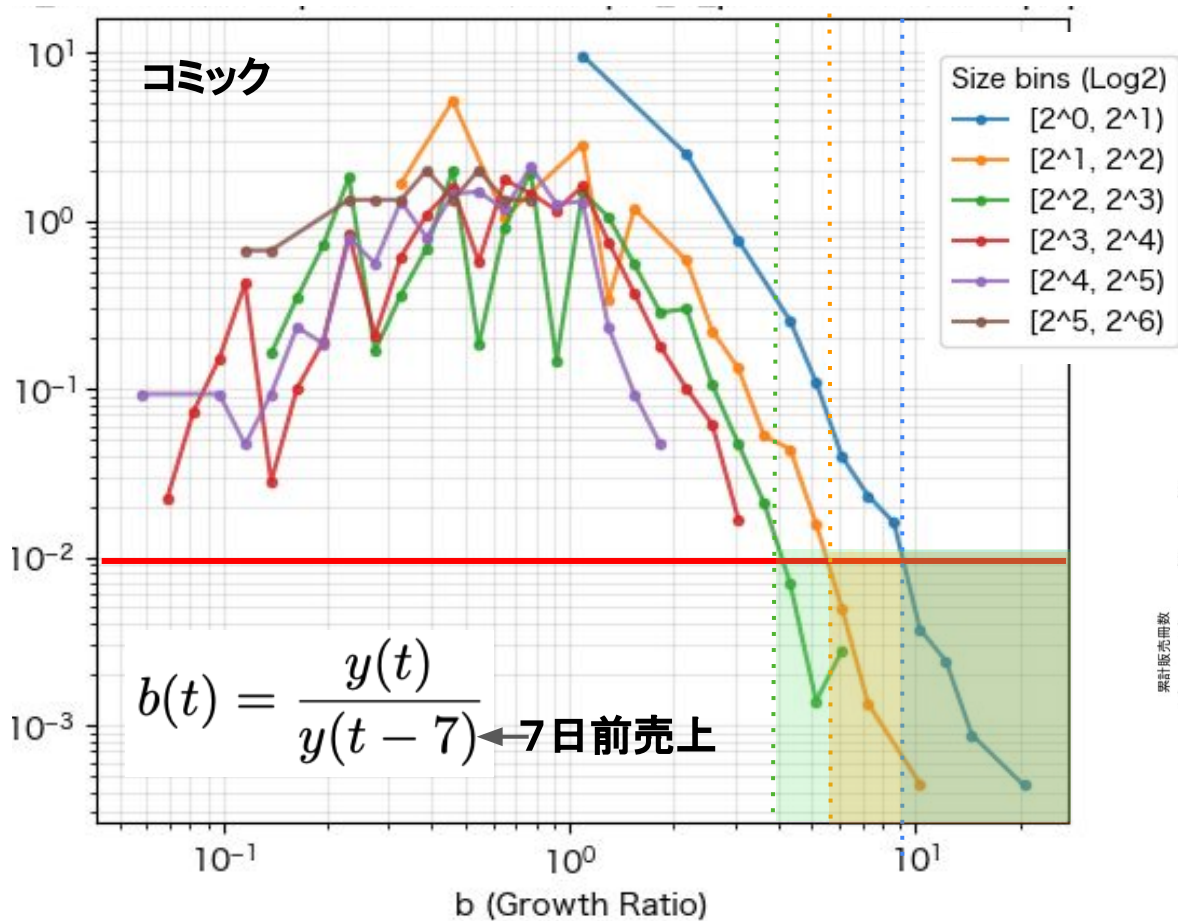
急上昇はどこか？

- ・仮に急上昇の考え方を「前日の売上と比べて今日の売上が何倍になったか」とする



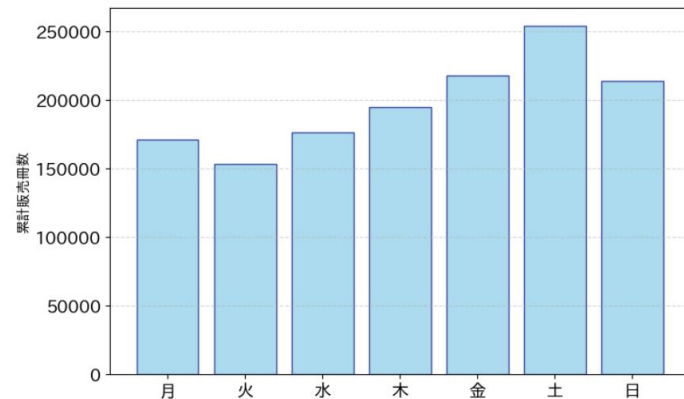
前日の売上規模を考慮した成長率の分布を確認

条件付き成長率の分布 (売上数でグループを分ける)

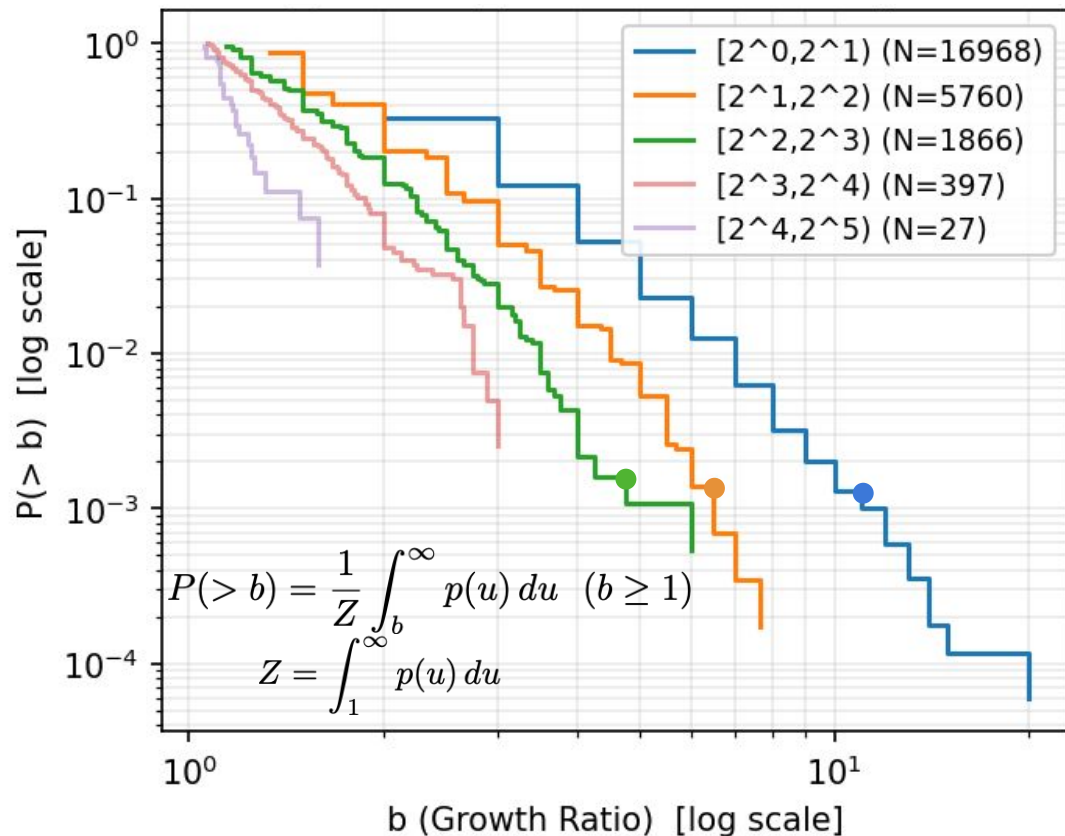


売上数大

売上数の曜日依存性



稀な急上昇イベントを持つタイトル



コミックから
確率 $P(>b) < 1/800$ の以下の稀な
成長率を記録するタイトルを選出。

大分類	タイトル数
コミック	44

④ ピーク後の減衰のモデル化 (指数orべき)

4つのモデル

- ① 指数関数モデル(定数項あり)³⁾

$$y(t) = A \exp(-bt) + C$$

- ② 指数関数モデル(定数項なし)²⁾

$$y(t) = A \exp(-bt)$$

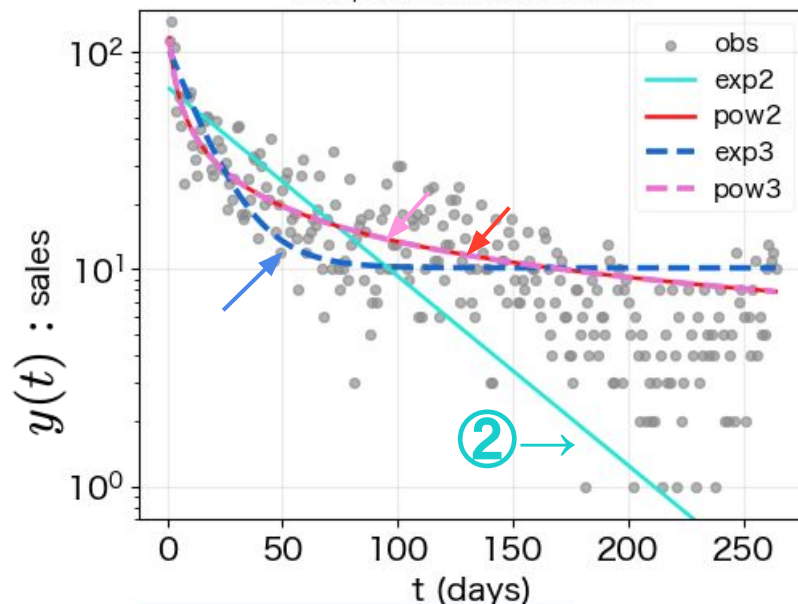
- ③ べき関数モデル(定数項あり)

$$y(t) = A (t + 1)^{-\alpha} + C$$

- ④ べき関数モデル(定数項なし)

$$y(t) = A (t + 1)^{-\alpha}$$

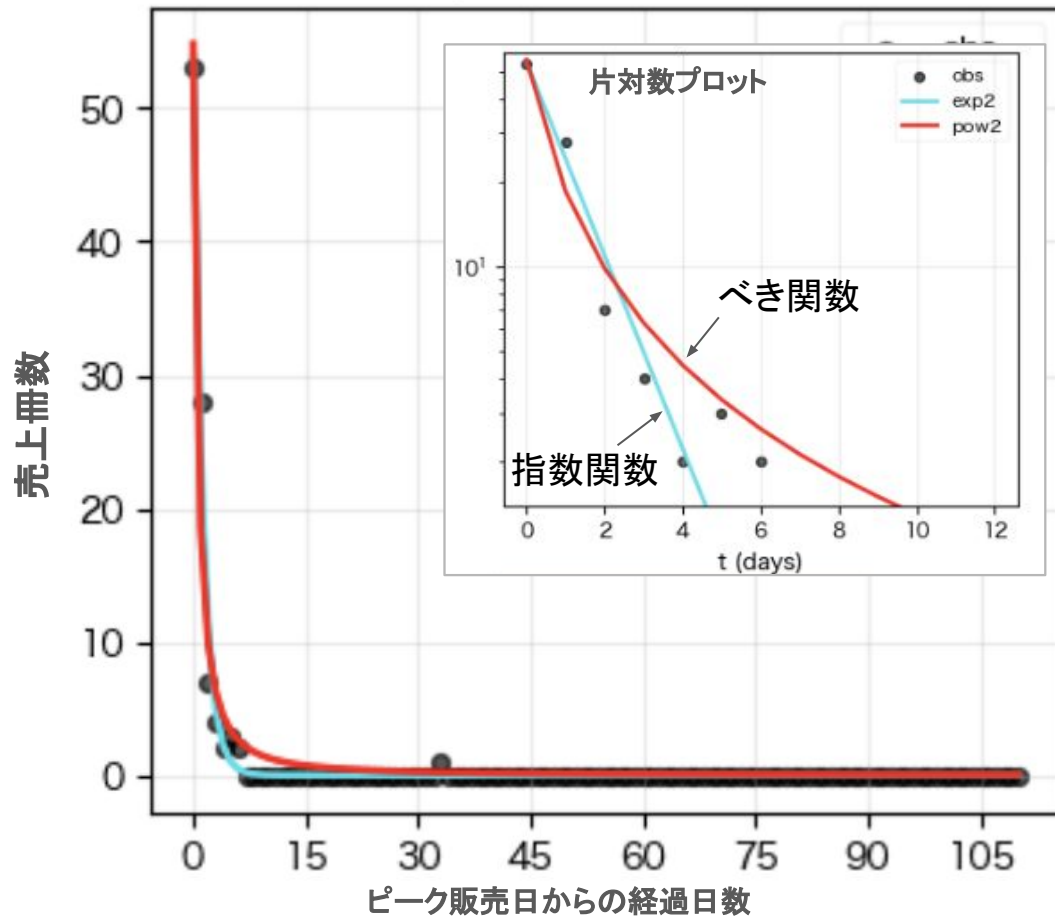
タイトルをフィッティング



・データ数が30以上($t \geq 30$)のみフィッティング

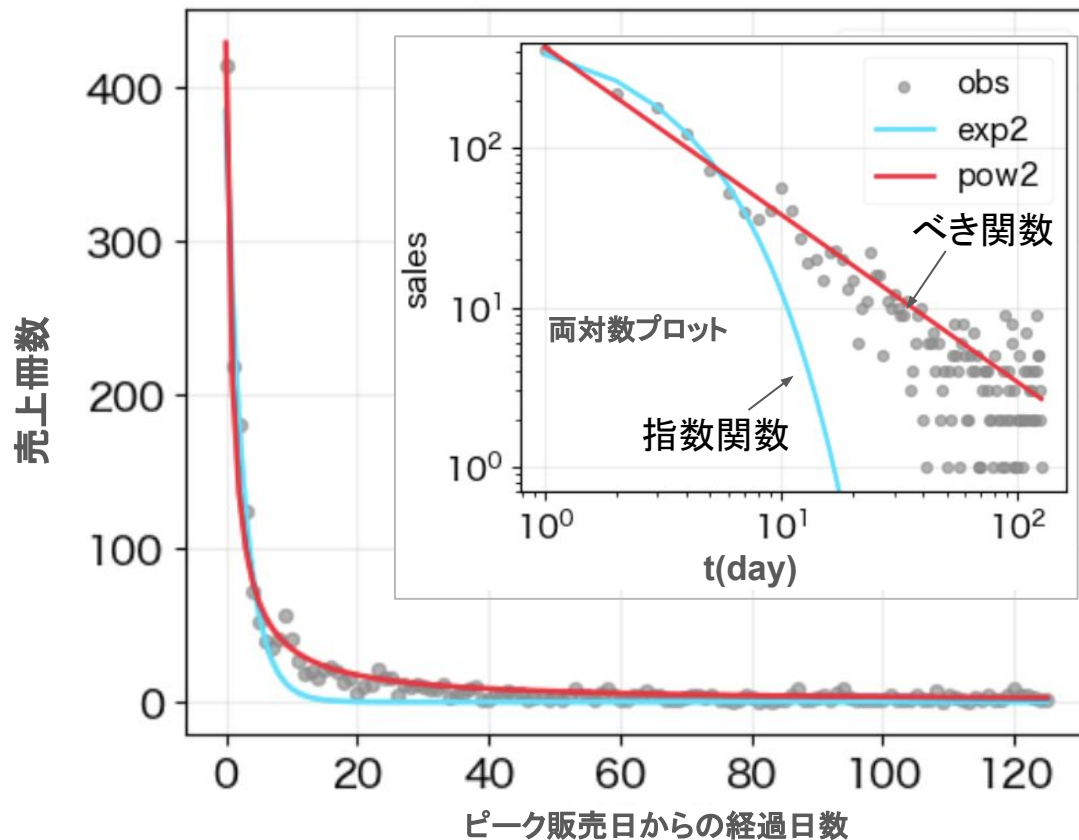
・①～④のモデルでフィッティング→BICを用いてモデル選択

指数関数モデル ($y(t) = A \exp(-bt)$) が選択されたサンプル



ピークから0冊付近までの減衰が早い

べき関数モデル ($y(t) = A(t + 1)^{-\alpha}$)が選択されたサンプル



ピーク直後は急激に減衰するが、その後はゆっくりと尾を引くように減衰(ロングテール)

書籍名: チェンソーマン17

急上昇検出 & (減衰予測)ダッシュボードを提案



Large (16-63冊) Scale: 2^4 - 2^6

ISBN: 978-4-10-354951-2 成瀬は天下を取りに行く +17.8倍 前: 37 冊 → 後: 660 冊	ISBN: 978-4-86680-936-6 書いてはいけない +8.6倍 前: 20 冊 → 後: 172 冊	ISBN: 978-4-10-354952-9 成瀬は信じた道をいく +5.6倍 前: 21 冊 → 後: 118 冊	
4 Mono Max (モノマックス)	26	→ 101	+3.9倍
5 山と食欲と私 18	18	→ 65	+3.6倍
6 Men's NONNO (メンズノンノ)	17	→ 60	+3.5倍
7 FAIRY TAIL 100 YE 17	25	→ 83	+3.3倍

まとめ

- ・売上規模によって成長率の分布は異なる
→システムティックな急上昇検出アルゴリズムを構築
- ・新発売や急上昇後の減衰は**指数型(35%)**と**べき型(65%)**
- ・ジャンル毎にべき指数の分布が異なる
- ・べき指数は最初の**3日**程度でおおよそ見積もることができる
- ・ピーク売上規模に応じて**べき指数を補正**し予測を改善できる
- ・店員さんが効率的に急上昇を見つけるためのダッシュボードを提案

課題

- ・同ジャンル内でのべき指数の違いの要因特定
- ・ダッシュボードに売上予測と在庫管理アラームの実装