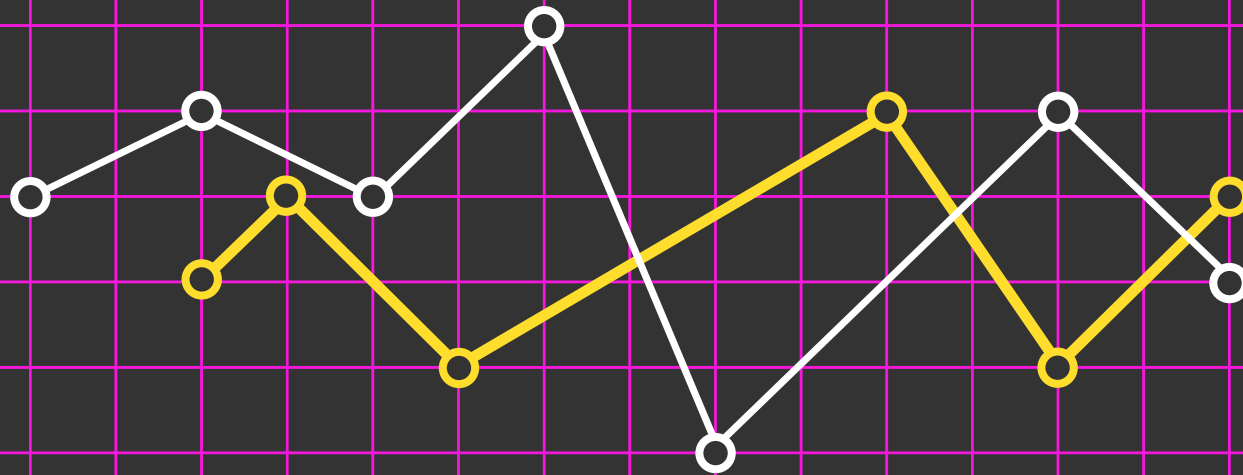


DANC



Команда T-NAME

1. Цель исследования

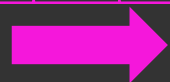
Определить, существует ли зависимость между педагогическим стажем (`experience_uni`) преподавателя и вероятностью того, что он получит звание «Лучший преподаватель» (`best_teacher`).

Гипотеза:

С увеличением стажа растет вероятность победы в конкурсе на "лучшего преподавателя", причем до определенного момента.

2. Механизм

Стаж увеличивается



Эффективнее учит



Голосует большое

Стаж увеличивается



Методы устаревают

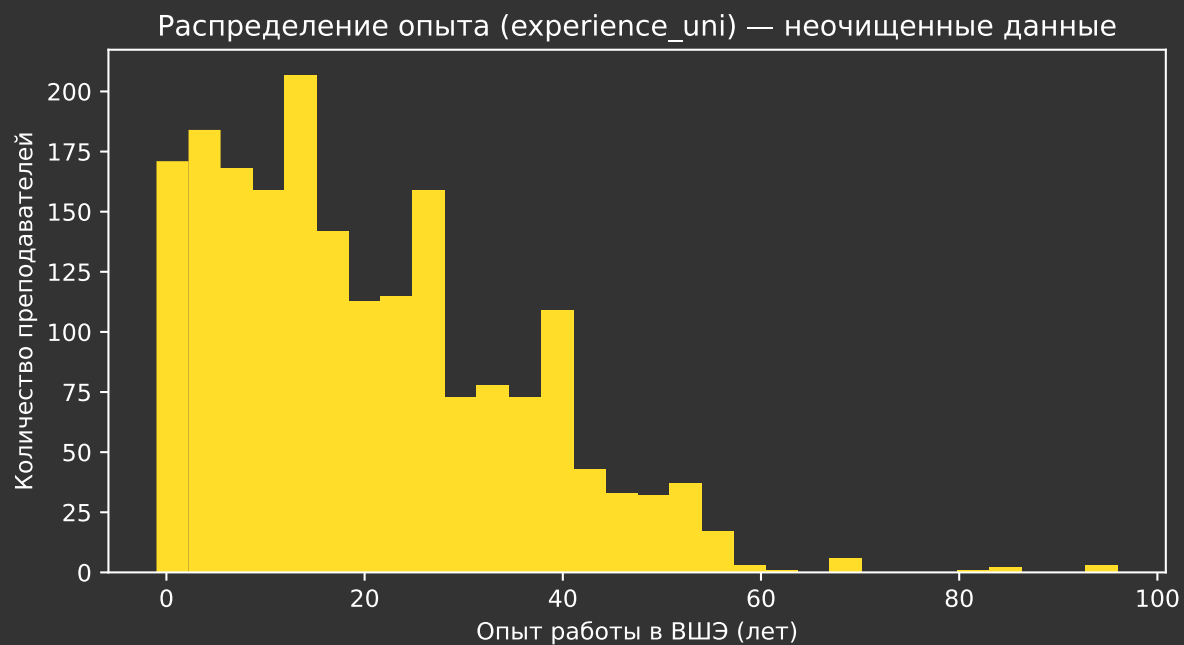


Голосует меньше

3. Разведочный анализ и очистка данных (EDA)

Перед анализом данных мы провели полный разведывательный анализ (EDA)

Брать значения больше 50 уже плохо, потому что дальше идут очень разрывные значения. И это испортит графики. Такие значения мы будем считать выбросами.



4. Разведочный анализ и очистка данных (EDA)

Разведка

В таблице было около 2487 строк

1) Переменная `experience_uni` содержала пропуски в 550 случаях ~22.4%

2) Среди непустых значений встречались. Таких записей оказалось менее 1%, но они критически искажают графики.

Очистка

1. Приведение `experience_uni` к числовому типу и `NaN` если такое отсутствует.

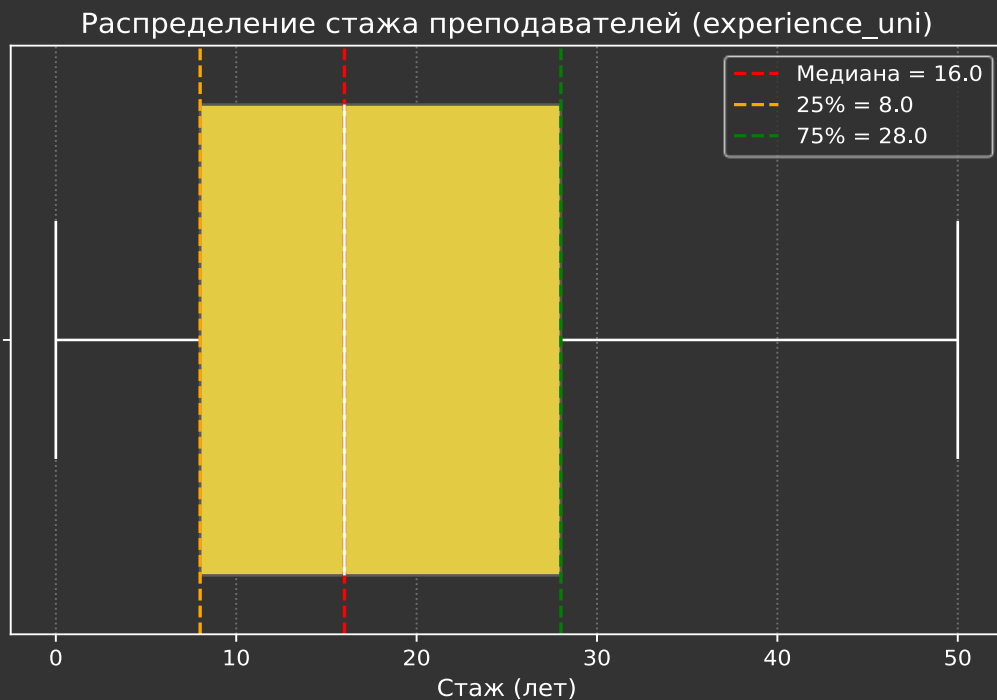
2. Исключение записей с `experience_uni < 0`
`experience_uni > 50`

Итог

После очистки, осталось 1840 строк корректных данных ~ 73% исходного массива.

На основании этих данных строится все дальнейшие расчёты

5. Распределение учителей по количеству опыта преподавания



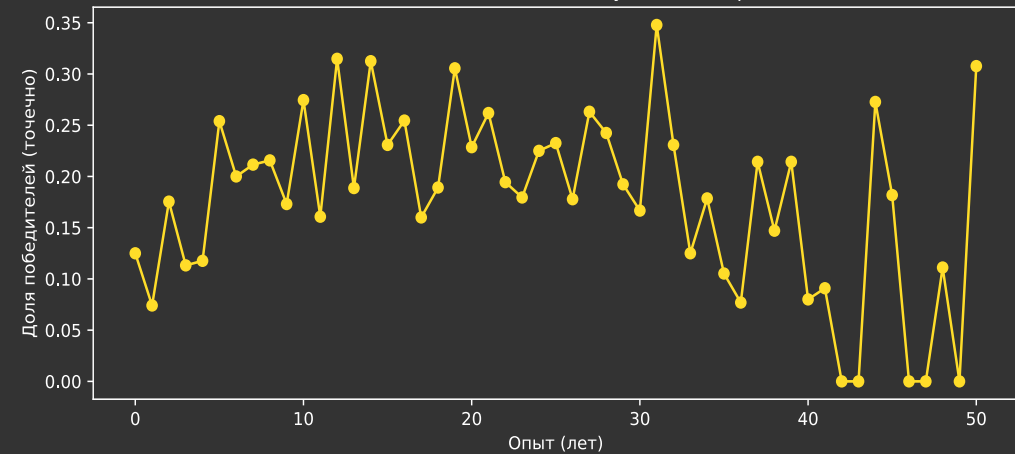
Из получившегося графика следует:

- 1) Большинство преподавателей имеют стаж до 30 лет
- 2) Пик — в районе 7–21 лет стажа
- 3) После 40 лет количество наблюдений резко снижается (менее 5% выборки). Это важно: такие редкие значения на графиках могут создавать «пики» или «нули», не отражающие реальную тенденцию

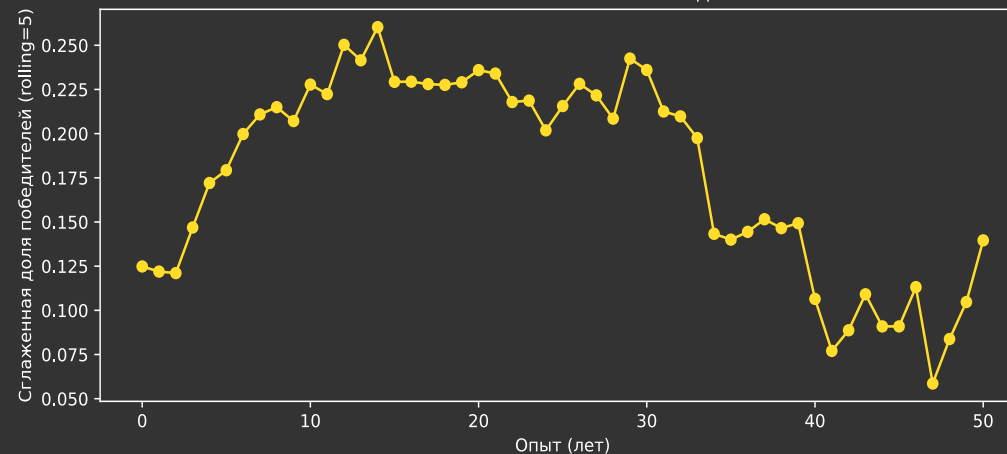
6. Сглаживание Данных

Для устранения выбросов было применено сглаживание данных методом "rolling window" это уберет шум, погрешности и покажет общий тренд.

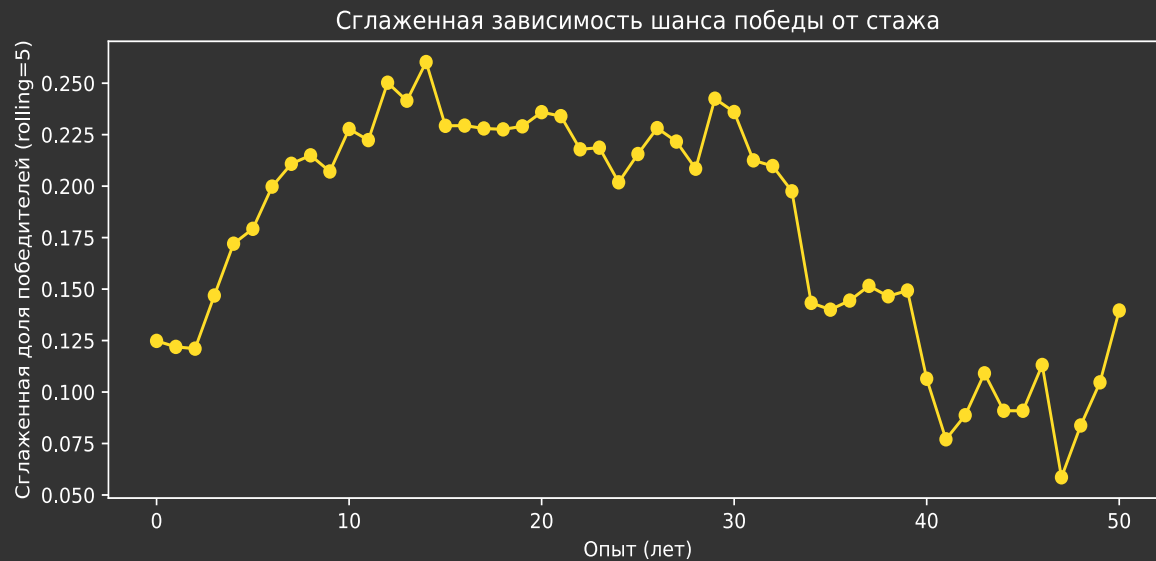
Точечная доля победителей по году стажа (сырые данные)



Сглаженная зависимость шанса победы от стажа



7. Визуализация и оценка зависимости



Выводы из графиков зависимости шанса победы от стажа:

1. Нулевой опыт (0–4 года): минимальные шансы, около 2–3%.
2. Плавный рост до 10–19 лет — пик вероятности около 18–26% победителей.
3. После 25–30 лет — устойчивое снижение до ~10%.
4. После 30 резкое снижение. На это влияет малое количество данных в выборке для этого диапазона.

8. Бинирование

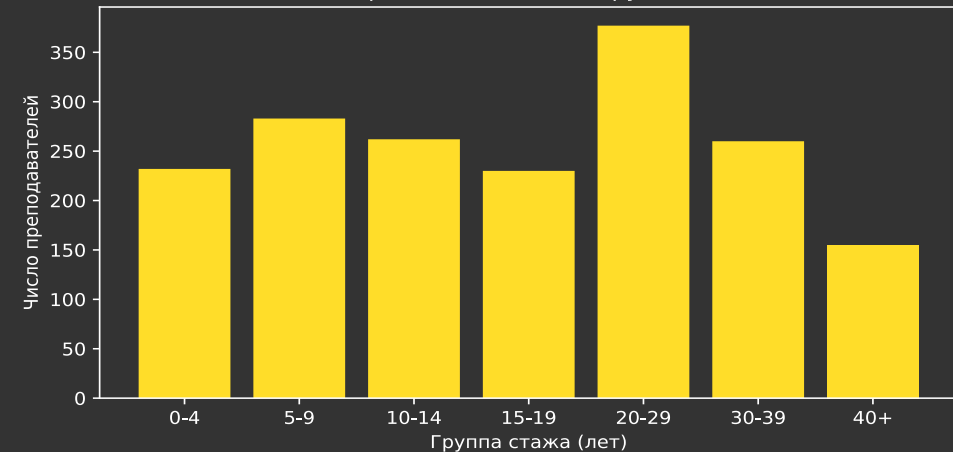
Для уменьшения шума полезно агрегировать наблюдения по интервалам. В отчёте используется следующая разбивка: 0–4, 5–9, 10–14, 15–19, 20–29, 30–39, 40+ лет. Для каждой группы вычисляются:

- абсолютное число преподавателей;
- доля победителей;
- доля от всей выборки

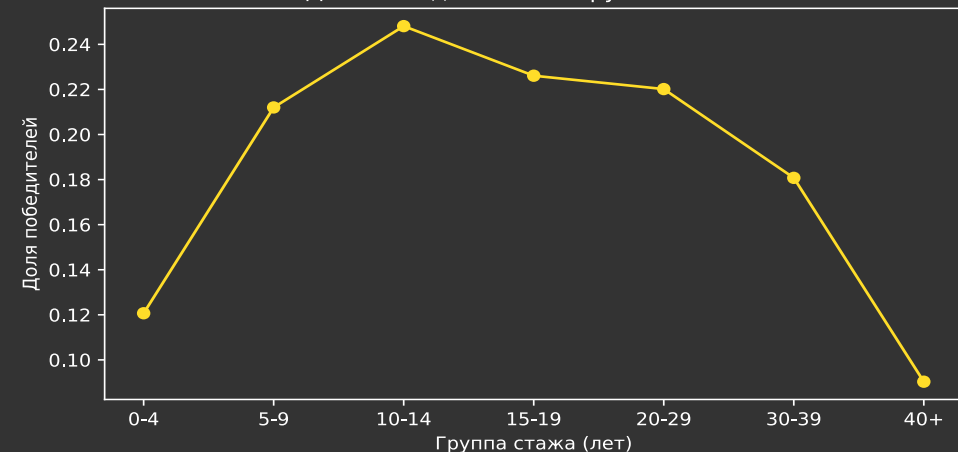
Интервалы удобны для формулирования политик

Пик доли победителей приходится на диапазон 10–19 лет, где вероятность почти в 4 раза выше, чем у новичков (20% против 4–5%). После 30 лет наблюдается спад.

Число преподавателей по группам стажа



Доля победителей по группам стажа



9. Сравнительный анализ по факультетам



Вывод: зависимость наиболее устойчива на гуманитарном факультете, менее выражена на экономическом и математическом.

Математика — фундаментальная наука, где ценится опыт и глубокие знания. Часто на таких факультетах остаются работать преподаватели, которые уже много лет преподают, потому что:

1. сложнее заменить опытных преподавателей новыми;
2. студенты и коллеги ценят стабильность и наставничество.

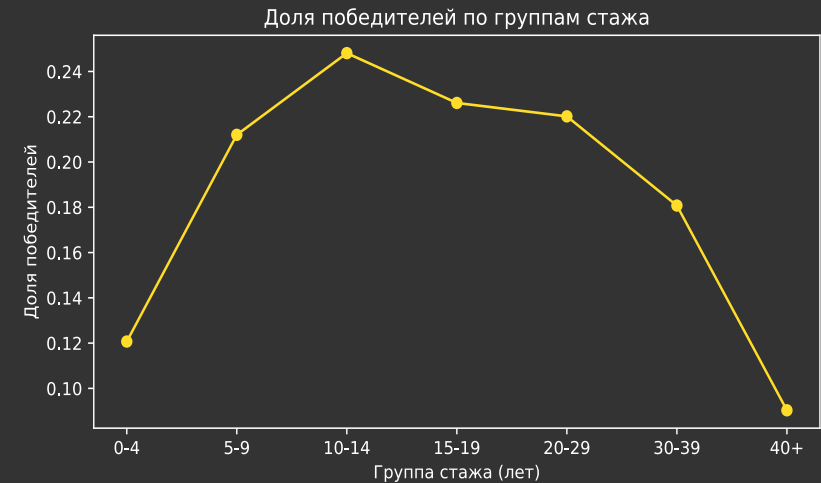


10. T-TEST

T-тест — это статистический метод, используемый для проверки гипотез о разнице между средними значениями.

Краткое описание:

Определить, существует ли статистически значимая разница между средними значениями двух групп (или между средним и известным значением)



0-4 vs 5-29

t отрицательный, потому что шанс у 0–4 меньше, чем у 5–29. $p \approx 0.00016 \rightarrow$ очень малое, значит различие статистически значимо.

Вывод: преподаватели со стажем 5–29 лет имеют значимо больше шансов стать победителем, чем 0–4.

0-4 vs 30+

t отрицательный, но $p \approx 0.35 \rightarrow$ не статистически значимо. Вывод: шансы 0–4 и 30+ не отличаются статистически.

5-29 vs 30+

t положительный, $p \approx 0.00032 \rightarrow$ очень малое, различие значимо. Вывод: преподаватели со стажем 5–29 лет имеют значимо больше шансов стать победителем, чем 30+.

11. Выводы

- Наблюдается выраженная зависимость между стажем и вероятностью получения звания: минимальные шансы у преподавателей с опытом 0–4 года, максимальные — в среднем стаже (~10–19 лет), последующее снижение для преподавателей со стажем >30 лет.
- Эффект выражен неодинаково по факультетам: наиболее заметен на human, менее — на еcom, и нестабилен на math из-за специфики факультета.

Рекомендации для политики ВШЭ

1. Разработать программы поддержки молодых преподавателей (0–4 года) — менторство, методические тренинги, отдельные поощрения.
2. Пересмотреть критерии поощрений для ветеранов, чтобы корректно учитывать их вклад (научный, административный).
3. Включать стаж как фактор в многомерные модели оценки деятельности, но не рассматривать его как единственный критерий.