

Отношение студентов Приамурского государственного университета имени Шолом-Алейхема к спортивной деятельности

Ересь Артём Владимирович

Приамурский государственный университет имени Шолом-Алейхема

Студент

Баженов Руслан Иванович

Приамурский государственный университет имени Шолом-Алейхема

к.п.н., доцент, зав. кафедрой информационных систем, математики и правовой информатики

Аннотация

В данной работе рассмотрен один из методов интеллектуального анализа данных - метод корреляционного анализа. На примере данных из результатов анкетирования среди студентов Приамурского государственного университета имени Шолом-Алейхема был вычислен коэффициент корреляции Спирмена. На основе него проведена оценка связи спорта и психологии.

Ключевые слова: язык программирования R, RStudio, статистика, корреляция, Спирмен.

Relation of students of the Sholom-Aleichem Priamursky State University to sporting activities

Yeres Artem Vladimirovich

Sholom-Aleichem Priamursky State University

Student

Bazhenov Ruslan Ivanovich

Sholom-Aleichem Priamursky State University

Candidate of pedagogical sciences, associate professor, Head of the Department of Information Systems, Mathematics and Law Informatics

Abstract

In this operation one of methods of data mining - a method of correlation analysis is considered. On the example of data from results of questioning among students of the Sholom-Aleichem Priamursky State University Spearman correlation coefficient was calculated. On the basis of it assessment of communication of sport and psychology is carried out.

Keywords: programming language R, Studio, statistics, correlation, Spearman.

В современном мире спорт является неотъемлемой частью жизни многих людей. Каждый из нас хоть раз посетил тренировку, либо присутствовал на занятиях физической культуры в школе, но не всякий знает о тонкостях этих дисциплин.

На самом деле, именно спорт оказывает активное влияние на становление многосторонне развитой личности. Оказывается спорт может быть не только средством укрепления здоровья или формой проведения свободного времени, но существенно влиять и на другие стороны человеческой жизни: на трудовую деятельность, нравственные и интеллектуальные качества, на формирование личности человека. Также физическая культура и спорт влияют на нравственно-этическое воспитание, процесс самопознания и самосовершенствования, на формирование воли.

Одним из основополагающих элементов формирования развития физической культуры в современном обществе является осознанная направленность личности. Ведь именно она определяет поведение личности, его отношение к себе и к окружающим. Существует понятие гармоничного развития человека, в которое помимо антропометрических данных, которые во многом определяются наследственными факторами, следует также включать единство и согласованность «нетелесных», присущих только человеку качеств (воля, нравственность, эстетичность, интеллект).

Следует отметить, что активный интерес к спорту проявляется у человека в школьные и студенческие годы. Именно в этом возрасте происходит становление личности, укрепление личностных качеств и уровня самооценки, выбор жизненных ориентиров.

Целью данного исследования является анализ зависимости результатов анкетирования по вопросу отношения студентов к спорту с результатами выявленного уровня психологического развития личности. Это позволит выявить зависимости между спортивной деятельностью и психологическим аспектом, показать тенденции интересов современной молодежи.

В нашей работе будет использован статистический метод, предназначенный для выявления существования зависимости между двумя и более случайными величинами - метод корреляционного анализа. В частности для наших данных будет использован корреляционный анализ Спирмена, что позволит более четко увидеть связи между переменными и показать зависимости.

Тема данного исследования достаточно актуальна в настоящее время. Рассмотрим существующие работы. М.И. Калугина и С.В. Бегичева обозначили современные возможности среды R, продемонстрировали способы визуализации итогов вычислений [1]. Ученые О.В. Прокофьев, И.Ю. Семочкина использовали для обработки математических данных программу RStudio, что позволило показать её положительные стороны [2]. Ф.М. Курилов с помощью языка R в своей статье провел эконометрическое моделирование и обеспечил точность расчетов [3]. С.С. Певченко выделил основные методы интеллектуального анализа данных, провел их полную классификацию, на основе которой можно подобрать под каждый тип

данных свой метод [4]. И.В. Зайцева, Т.Р. Лысак и Т.А. Горюнова охарактеризовали применение ИТ-технологий для статистический вычислений [5]. В книге группы психологов Н. П. Фетискина, В. В. Козлова и Г. М. Мануйлова определены методики психологического анализа личности, проведена классификация основных типов развития человека [6]. На иноязычных источниках так же достаточно работ, которые касаются тематики нашей статьи [7-8]. Множество интернет-ресурсов позволяют подробно разобраться в корреляционном анализе Спирмена и в применении языка R [9-13].

Коэффициент ранговой корреляции Спирмена используется для выявления и оценки связи между двумя рядами сопоставляемых количественных показателей. В том случае, если ранги показателей, упорядоченных по степени возрастания или убывания, в большинстве случаев совпадают, делается вывод о наличии прямой корреляционной связи. Если ранги показателей имеют противоположную направленность, то говорят об обратной связи между показателями.

Коэффициент корреляции Спирмена обладает следующими свойствами:

- Коэффициент корреляции может принимать значения от минус единицы до единицы, причем при $r_s=1$ имеет место строго прямая связь, а при $r_s=-1$ – строго обратная связь.
- Если коэффициент корреляции отрицательный, то имеет место обратная связь, если положительный, то – прямая связь.
- Если коэффициент корреляции равен нулю, то связь между величинами практически отсутствует.
- Чем ближе модуль коэффициента корреляции к единице, тем более сильной является связь между измеряемыми величинами.

Для существенного упрощения и сокращения времени работы применим язык программирования R и соответствующую для него среду "RStudio". Этот программный продукт имеет множество особенностей, что делает работу с R очень удобной. R поддерживает различные методы статистического анализа, а так же имеет возможность расширяться за счет подключения дополнительных пакетов. Пакеты представляют собой библиотеки для работы специфических функций или специальных областей применения.

Перед началом работы в программе рассмотрим теоретические основы нахождения коэффициента корреляции Спирмена.

Расчет коэффициента ранговой корреляции Спирмена включает следующие этапы:

- 1) Сопоставить каждому из признаков их порядковый номер (ранг) по возрастанию или убыванию.
- 2) Определить разности рангов каждой пары сопоставляемых значений (d).
- 3) Возвести в квадрат каждую разность и суммировать полученные результаты.

4) Вычислить коэффициент корреляции рангов по формуле:

$$\rho = 1 - \frac{6 \cdot \sum d^2}{n(n^2 - 1)}$$

При использовании коэффициента ранговой корреляции условно оценивают тесноту связи между признаками, считая значения коэффициента равные 0,3 и менее - показателями слабой тесноты связи; значения более 0,4, но менее 0,7 - показателями умеренной тесноты связи, а значения 0,7 и более - показателями высокой тесноты связи.

Теперь приступим к использованию самой программы «RStudio» для нахождения коэффициента ранговой корреляции наших данных.

Для начала добавим наш файл `oprosG.csv`, для этого используем следующую команду:

```
> s <- read.csv("oprosG.csv", header = TRUE, sep = ";")
```

Рис.1. Команда для добавления данных в программу RStudio

Далее необходимо найти коэффициент корреляции Спирмена. Воспользуемся следующим методом: заменой ответов, где наихудший результат в вопросе принимается за ноль, и далее по нарастающей. В заданном файле `oprosG.csv` с помощью замены проведем операцию по смене данных на определенные значения, где ответы принимаются за значения в диапазоне от 0 до 5 по возрастающей. После этого просуммирует каждую строку ответов и получим новый столбик `"sum.FIZ"`.

Вводим следующую команду:

```
> cor.test(S$sum.FIZ, S$sum.PSIX, method = "spearman")
```

Рис.2. Команда для вычисления корреляции Спирмена

В получаем результат:

```

      spearman's rank correlation rho
data:  S$sum.FIZ and S$sum.PSIX
S = 125050, p-value = 0.01226
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.249606

```

Рис.3. Результат вычисления корреляции Спирмена

В результате получим коэффициент корреляции 0,24. Это показывает наличие низкой связи между двумя переменными, стоит отметить, что всё же спортивную деятельность возможно сопоставлять с уровнем саморазвития.

Для более наглядного вывода выведем на экран график зависимостей наших переменных. Добавим в систему библиотеку `ggplot2` и введем следующую команду в консоль:

```
ggplot(D,aes(sum.FIZ, sum.PSIX))+geom_line()+geom_point(size=2)+geom_smooth(method = "lm")
```

Рис.4. Команда вывода графика зависимости переменных

И получим следующий график (рис.5).

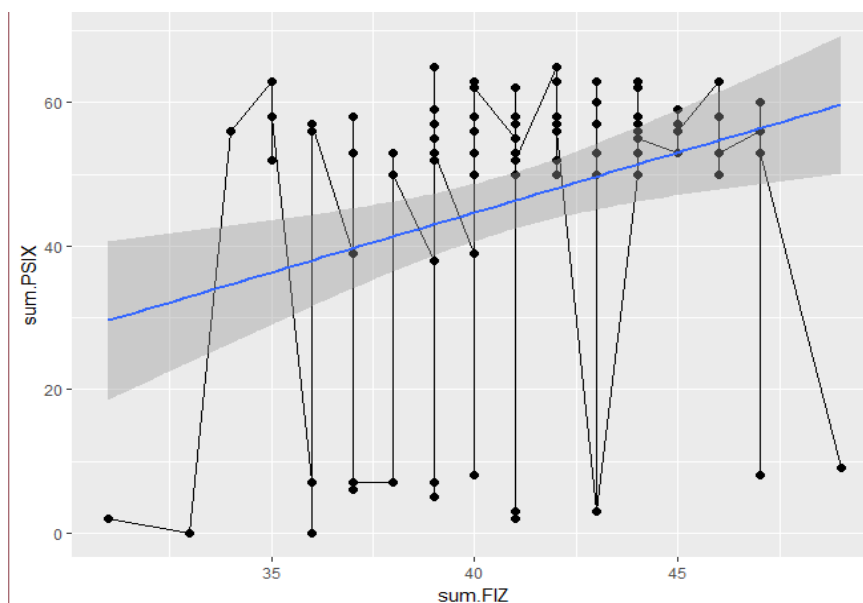


Рис.5. График зависимости переменных

На основе графика можно сделать выводы, что наибольшие значения психологической суммы находится у физической суммы в диапазоне от 30 до 50 единиц.

Рассмотрим круговые диаграммы и на основе них проанализируем интересы участников опроса. Первые 3 вопроса анкеты проверяют пол, курс и возрастную группу. По половой принадлежности участниками являются 50% женского пола и 50% мужского. По возрасту большинство участников принадлежит к группе от 21 до 25 лет, это доказывает, что в вопросе курса обучения высокие проценты имеют именно курсы студентов этого возраста.

В результате опроса был сделан вывод что большинство студентов, а именно 74% активно проявляют интерес к спорту и занимаются им.

Занимаетесь ли Вы спортом?

100 ответов

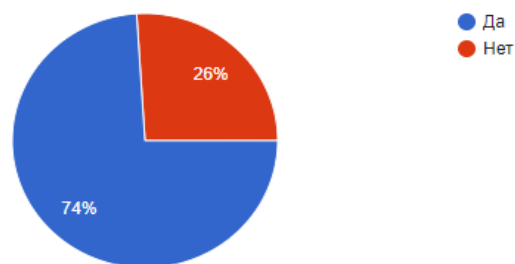


Рис.6. Диаграмма вопроса "Занимаетесь ли Вы спортом?"

По вопросу утренней зарядке можно сделать достаточно неожиданные выводы. Не смотря на то, что современная молодежь имеет склонность к поздним утренним подъемам, все же есть 11% ежедневно делающие зарядку. Это говорит о уровня самоорганизации и режима дня у некоторых студентов. Что в свою очередь является положительным явлением среди молодежи.

Делаете ли Вы утреннюю зарядку?

100 ответов



Рис.7. Диаграмма вопроса "Делаете ли Вы утреннюю зарядку?"

В вопросе проведения свободного времени следует обратить внимание на такое времяпровождение как постоянное нахождение перед компьютером. Это показывает, что современные студенты чаще выбирают компьютер, чем прогулки и свежий воздух, что свидетельствует о низком уровне интереса к активной деятельности.

Как Вы предпочитаете отдыхать?

99 ответов

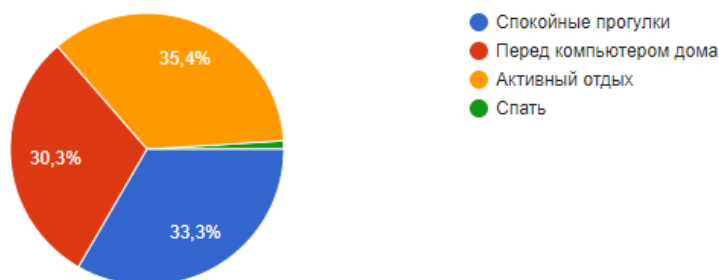


Рис.8. Диаграмма вопроса "Как Вы предпочитаете отдыхать?"

Вопрос о частоте занятия спортом отражает положительные результаты в плане того, что все же 47,5% несколько раз в неделю проводят свое время со спортом. Однако имеется отрицательный аспект - 14,1% совершенно не интересуются спортом. Несомненно этот факт несет в себе возможные проблемы со здоровьем, отсутствие физической подготовки и силы духа.

Как часто Вы занимаетесь спортом?

99 ответов

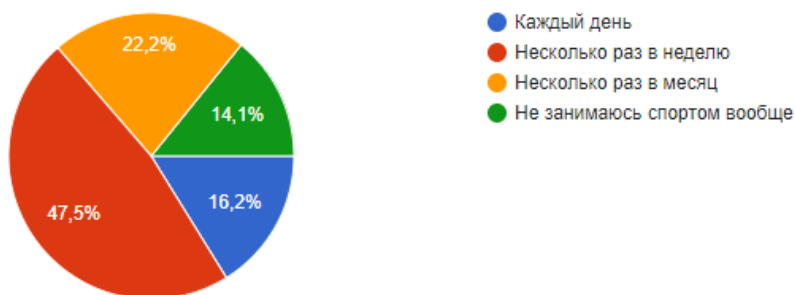


Рис.9. Диаграмма вопроса "Как часто Вы занимаетесь спортом?"

Следующая диаграмма показывает, что все-таки современных студентов к спорту приводит собственное желание и интерес. Это подразумевает осознанный выбор и доказывает значимость и ценность спортивных занятий для каждого. Достаточное влияние оказывает и дружба, что говорит о сплоченности и общих интересах групп студентов.

Как спорт появился в Вашей жизни?

97 ответов



Рис.10. Диаграмма вопроса "Как спорт появился в Вашей жизни?"

Результаты следующего вопроса отражают проблемы студентов в плане собственной неорганизованности и состоянии здоровья, когда именно эти факторы мешают человеку заниматься спортом. Так же отметим, что времени у студента хватает далеко не всегда и поэтому основной проблемой является нехватка времени на спорт. Из этого стоит сделать выводы о высокой загруженности студентов учебной и другой деятельностью.

Что мешает Вам заниматься спортом?

94 ответа

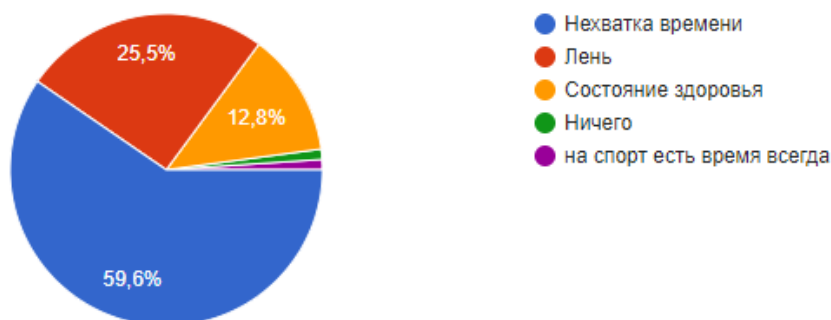


Рис.11. Диаграмма вопроса "Что мешает Вам заниматься спортом?"

Вредные привычки оказывают негативное влияние на спортивные занятия. Результаты по этому вопросу оказались хорошими и подавляющее большинство студентов не курят и не употребляют алкогольные напитки. Это говорит о наличии других интересов у людей, вместо пагубных влияний на свое здоровье.

Таким образом, в ходе работы был выявлен коэффициент корреляции достаточно низкого уровня, отражающий малую зависимость между переменными. Однако он не настолько приближен к минимальному и все же связь имеется. Для получения более высоких результатов возможно использование более четкой выборки участников, что существенно поднимет точность вычислений. Программа RStudio существенно сократила длительность вычислений, что говорит о ее простоте и надежности.

Библиографический список

1. Калугина М.И., Бегичева С.В. Современные возможности визуализации результатов исследований в среде R // ВІ-технологии корпоративные информационные системы в оптимизации бизнес-процессов. 2016. С. 51-55.
2. Прокофьев О.В., Семочкина И.Ю. Применение языка R и среды RStudio для математической обработки данных // Современные информационные технологии. 2017. № 25 (25). С. 47-51.
3. Курилов Ф. М. Использование языка R для эконометрического моделирования и обеспечения расчетов // Проблемы и перспективы экономики и управления. 2014. С. 15-17.
4. Певченко С. С. Методы интеллектуального анализа данных // Молодой ученый. 2015. №13. С. 167-169.
5. Зайцева И.В., Лысак Т.Р., Горюнова Т.А. Применение IT-технологий в статистике. // Учет, анализ, аудит и статистика: исследование тенденций и перспективы развития. 2016. С. 408-413.
6. Фетискин Н.П., Козлов В. В., Мануйлов Г. М. Социально-психологическая

- диагностика развития личности и малых групп // Изд-во Института Психотерапии. 2002.
7. Bunn A., Korpela M. R: A language and environment for statistical computing. 2013.
 8. Packianather M.S., Davies A., Harraden S., Soman S., White J. Data Mining Techniques Applied to a Manufacturing SME. Т. 62. 2017. С. 123-128.
 9. RStudio // Wikipedia URL: <https://ru.wikipedia.org/wiki/RStudio> (дата обращения 17.04.18)
 10. Коэффициент Спирмена // URL: <http://www.medstatistic.ru/theory/spirmen.html> (дата обращения: 18.04.2018).
 11. Корреляция // Wikipedia URL: <https://ru.wikipedia.org/wiki/Корреляция> (дата обращения 19.04.18).
 12. R: анализ и визуализация данных // URL: <http://r-analytics.blogspot.ru/p/rstudio.html#.Wurk5GtWr58> (дата обращения: 20.04.2018).
 13. Коэффициент корреляции Спирмена. Коэффициент ранговой корреляции Спирмена // URL: <http://fb.ru/article/149593/koeffitsient-korrelyatsii-spirmena-koeffitsient-rangovoy-korrelyatsii-spirmena> (дата обращения: 20.04.2018).