

Номинация 1: «Учимся собирать статистические данные»

Тема исследования: «Спорт и статистика. Сбор статистических данных из открытых источников и поиск закономерностей»

Суть статистики состоит в том, что с её помощью выявляются закономерности в окружающем нас мире. В основном, это касается массовых, то есть часто повторяющихся явлений. Если вас заинтересовало какое-то явление, и вы хотите понять, что, собственно, происходит, вы можете собрать информацию об этом явлении – провести наблюдение или поставить эксперимент, а затем попытаться выявить закономерности и объяснить их.

Самая трудоемкая часть этой работы – сбор первичных данных. Чтобы провести опрос или поставить эксперимент, надо затратить много сил и времени, для этого могут понадобиться определенные ресурсы или оборудование. Это не всегда и не всем доступно. Но иногда для поиска статистических закономерностей можно воспользоваться уже готовыми первичными данными, которые находятся в открытых источниках, например, в книгах, справочниках или в интернете. При этом надо уважать труд тех, кто собрал для вас эти первичные данные, и всегда указывать, откуда вы взяли данные, кто их автор и владелец.

Конечно, надо пользоваться только надежными источниками. Ваша задача – установить истинную закономерность, а не создать «жареную сенсацию», как говорят некоторые недобросовестные исследователи или журналисты. Если у вас возникли сомнения в достоверности найденных вами данных, лучше не используйте их, а поищите другой источник, более надежный.

Приведем пример статистической закономерности, которую удалось обнаружить, исследуя статистические данные. Этот пример связан со спортом.

В 1980-х годах исследователи обнаружили в некоторых ведущих спортивных лигах мира следующий интересный эффект. Если посмотреть на распределение дней рождений спортсменов, выступающих в лиге, по месяцам

года, то оказывается, что чем ближе месяц к концу года, тем меньше число спортсменов в лиге родилось в этом месяце! При этом дни рождений в общей популяции населения стран распределены равномерно в течение года. Почему тогда спортсменов, родившихся в январе, больше, чем спортсменов, родившихся в декабре? Оказывается, что дело в механизме отбора спортсменов. Набор в детские спортивные школы часто производится по году рождения, в результате чего внутри одной группы конкурируют друг с другом дети, которые родились в течение одного календарного года. Ребенок, родившийся в январе, имеет преимущество в физическом развитии почти в 1 год перед тем, кто родился в декабре того же года. В детском возрасте разница в 1 год является очень ощутимой, поэтому, когда наступает время отбора для перехода на следующий год обучения в спортивной школе, дети, родившиеся ближе к началу года, чаще показывают себя лучше и выигрывают конкуренцию. Этот эффект назвали эффектом относительного возраста. Эффект часто является достаточно устойчивым и наблюдается не только в детском возрасте, но и во взрослых лигах.

В рамках выполнения конкурсного задания Вам предстоит сделать следующее.

1. Выберите любую конкурентную область человеческой деятельности, устройство которой вы понимаете и для которой в открытых источниках можно найти данные о днях рождений участников этой деятельности. Это может быть, например, профессиональная женская футбольная лига или парламент Монголии, члены совета директоров 100 ведущих компаний мира или Нобелевские лауреаты.

2. До того, как приступить к сбору и анализу данных, сформулируйте гипотезу: как, на Ваш взгляд, должны быть распределены дни рождений участников рассматриваемой Вами области деятельности? Можно ли ожидать, что имеет место эффект относительного возраста? Почему?

3. Соберите базу данных дней рождений участников области деятельности, которую Вы рассматриваете. Это должна быть достаточно

большая база данных, чтобы в ней, согласно закону больших чисел, могли проявиться закономерности.

4. Постройте распределение дней рождений в вашей базе данных. Есть ли видимые закономерности в этом распределении? Сравните наблюдаемые результаты и Вашу гипотезу. Подтвердилась ли гипотеза?

5. Опишите, как устроен конкурентный отбор в выбранной Вами области деятельности. Заложены ли в механизме конкуренции предпосылки для тех или иных статистических особенностей, в том числе эффекта относительного возраста? Дайте возможные обоснования полученных Вами результатов с учетом описанного механизма конкуренции.

6. Подготовьте конкурсную работу, оформите ее так, чтобы было понятно, кто именно выполнил работу, под чьим руководством, откуда взяты первичные данные, какие методы вы применяли для их поиска, сбора и обработки.