

Лабораторная работа

Обработка результатов опроса

В ходе проведения опроса по теме “ Компетенции и процесс обучения у студентов ИТ направлений ”, было получено 56 ответов, которые были скачаны с Яндекс Форм в виде сводной таблицы (оригинал представлен в Excel файле с обработкой результатов опроса).

[illegible]

С помощью обработки результатов необходимо было получить ответы на следующие вопросы:

1. Как распределяются студенты по опыту в программировании (сколько лет они занимаются программированием)?
2. Какой язык программирования наиболее часто применяется студентами для написания проектов и реализация сложных задач?
3. Какие концепции программирования наиболее известны студентам?
4. Какие шаблоны (паттерны) проектирования наиболее известны или применяются студентами?
5. Какой формат обучения является наиболее предпочтительным для студентов?
6. Коррелируют ли языки программирования, изучаемые на направлениях обучения и любимые языки программирования студентов?
7. Коррелируют ли выбираемые формы представления материала по лекционным занятиям и лабораторным занятиям?

Результаты были обработаны двумя способами: при помощи программы Excel и при помощи программы KNIME.

Рассмотрим обработку результатов для получения ответов на первые пять вопросов.

Обработка через Excel

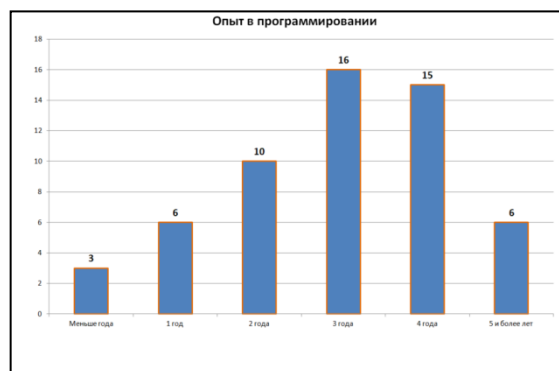
(Файл со всеми таблицами и диаграммами приложен в ответе)

1. Опыт в программировании

Таблица с количеством выборов сделанных в пользу каждого из вариантов (Лист “Опыт в программировании”):

	A	B
1	Опыт	Количество
2	Меньше года	3
3	1 год	6
4	2 года	10
5	3 года	16
6	4 года	15
7	5 и более лет	6

По данной таблице была построена диаграмма (Лист “Диаграмма_Опыт_в_прог...”):

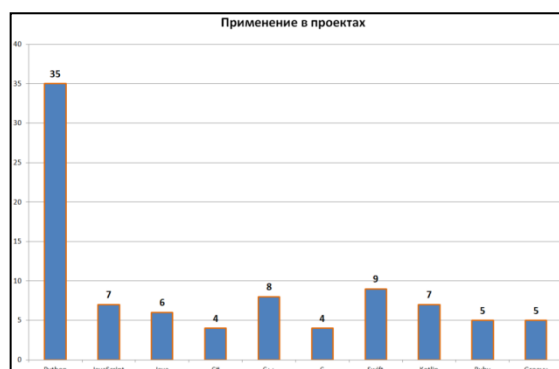


2. Применение в проектах

Таблица с количеством выборов сделанных в пользу каждого из вариантов (Лист “Применение в проектах”):

	A	B
1	Язык	Количество
2	Python	35
3	JavaScript	7
4	Java	6
5	C#	4
6	C++	8
7	C	4
8	Swift	9
9	Kotlin	7
10	Ruby	5
11	Groovy	5

По данной таблице была построена диаграмма (Лист “Диаграмма_Применение_в_про...”):

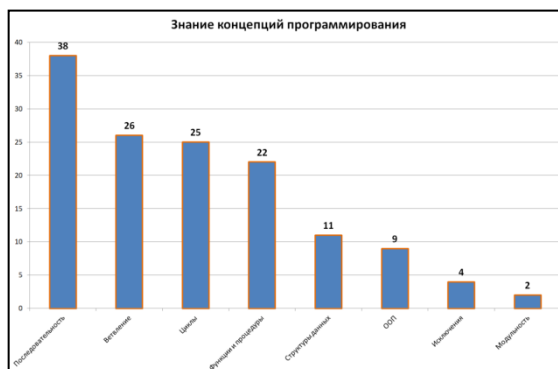


3. Знание концепций программирования

Таблица с количеством выборов сделанных в пользу каждого из вариантов (Лист “Концепции программирования”):

	А	В
1	Концепция	Количество
2	Последовательность	38
3	Ветвление	26
4	Циклы	25
5	Функции и процедуры	22
6	Структуры данных	11
7	ООП	9
8	Исключения	4
9	Модальность	2

По данной таблице была построена диаграмма (Лист “Диаграмма_Концепции_прог...”):

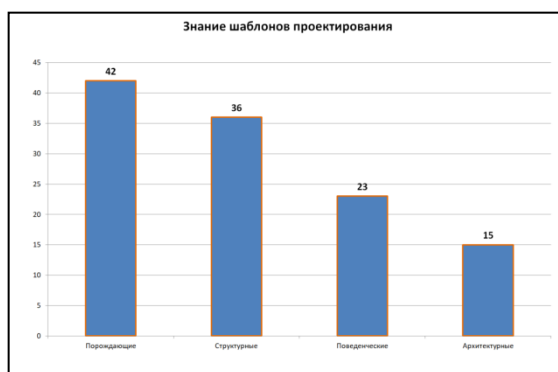


4. Знание шаблонов проектирования

Таблица с количеством выборов сделанных в пользу каждого из вариантов (Лист “Шаблоны проектирования ”):

	А	В
1	Паттерны	Количество
2	Порождающие	42
3	Структурные	36
4	Поведенческие	23
5	Архитектурные	15

По данной таблице была построена диаграмма (Лист “Диаграмма_Шаблоны_проект...”):

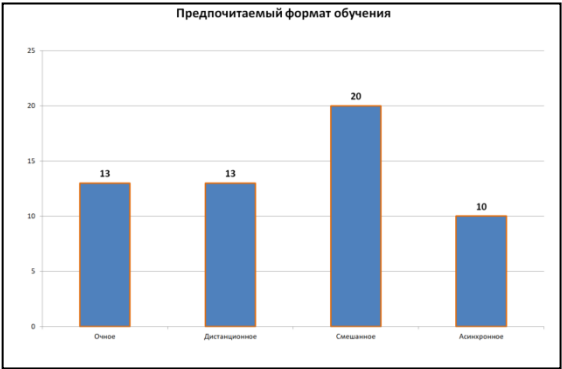


5. Предпочитаемый формат обучения

Таблица с количеством выборов сделанных в пользу каждого из вариантов (Лист “Формат обучения ”):

	А	В
1	Формат	Количество
2	Очное	13
3	Дистанционное	13
4	Смешанное	20
5	Асинхронное	10

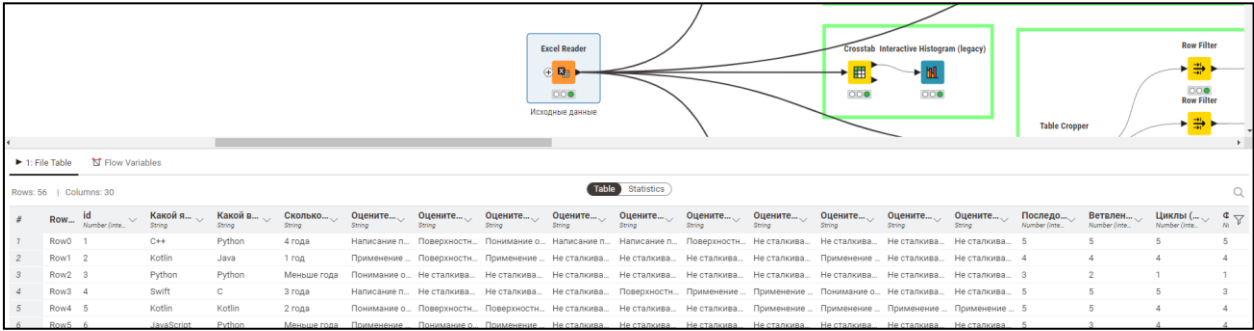
По данной таблице была построена диаграмма (Лист “Диаграмма_Формат_обучения”):



Обработка через KNIME

(Архив с разработанным с помощью KNIME проектом приложен в ответе)

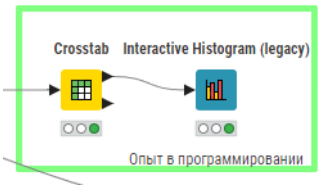
С помощью “Excel Reader” в проект были загружены данные из исходной таблицы с результатами опроса:



Затем была проведена обработка результатов для каждого вопроса и составлена общая схема.

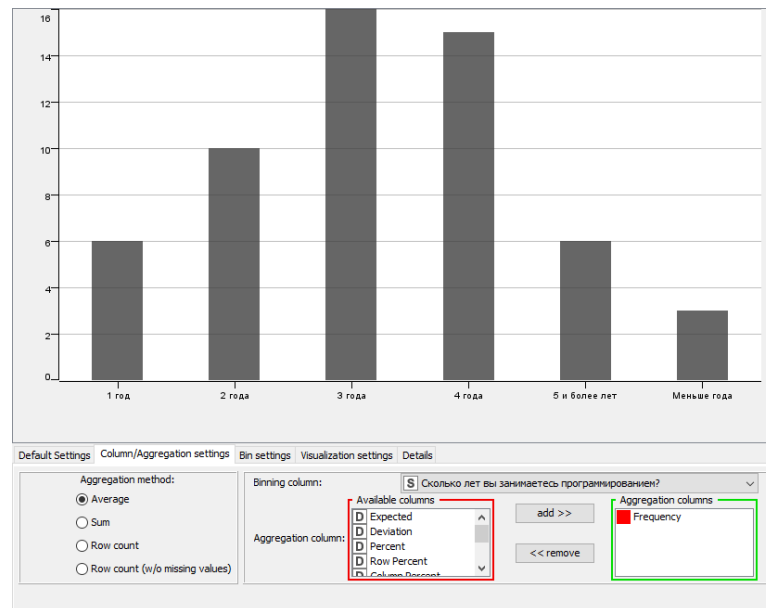
1. Опыт в программировании

За обработку данных результатов отвечает следующий фрагмент схемы (выбирает необходимые данные из исходной таблицы, а затем строит диаграмму):



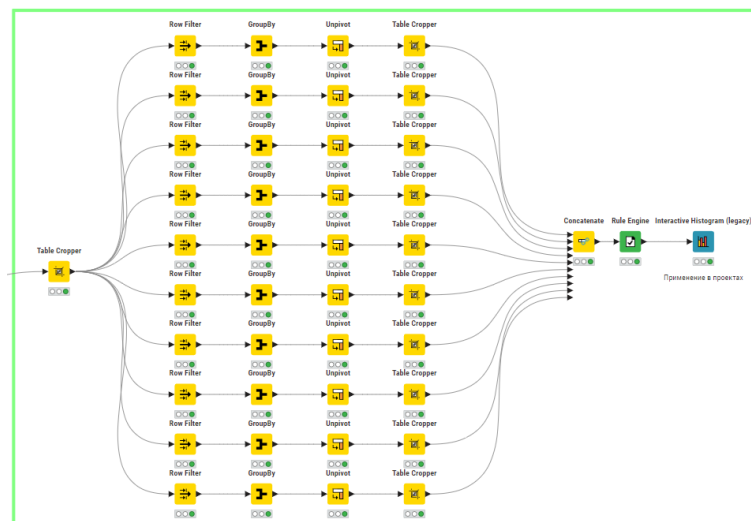
На выходе получаем таблицу со значениями (количеством выборов для каждого варианта) и диаграмму распределения:

#	Row...	Сколько л...	-	Frequency
	String	String		Number (double)
1	Row0	1 год	-	6
2	Row1	2 года	-	10
3	Row2	3 года	-	16
4	Row3	4 года	-	15
5	Row4	5 и более лет	-	6
6	Row5	Меньше года	-	3



2. Применение в проектах

За обработку данных результатов отвечает следующий фрагмент схемы (выбирает и обрабатывает необходимые данные из исходной таблицы, а затем строит диаграмму):



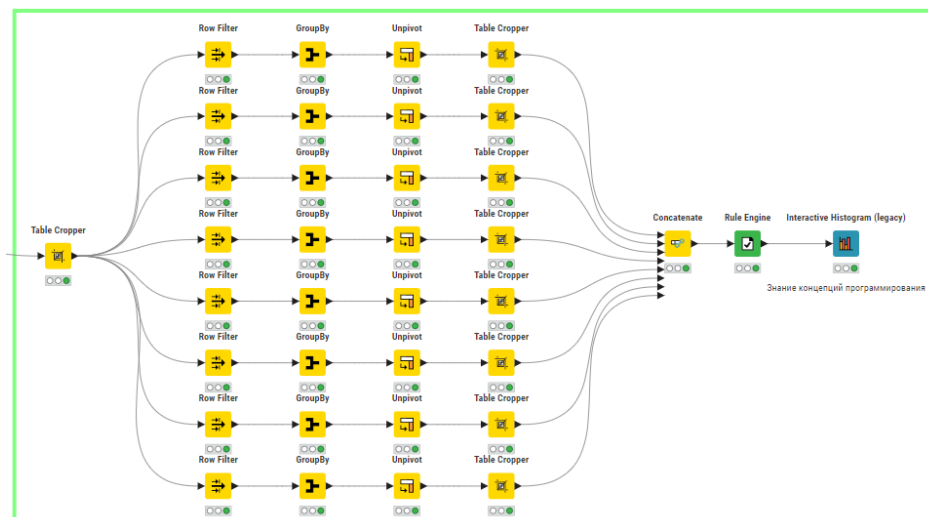
На выходе получаем таблицу со значениями (количеством выборов для каждого варианта) и диаграмму распределения:

#	Row...	ColumnNames String	ColumnValues Number (integer)
1	Row0	Python	35
2	Row...	JavaScript	7
3	Row...	Java	6
4	Row...	C#	4
5	Row...	C++	8
6	Row...	C	4
7	Row...	Swift	9
8	Row...	Kotlin	7
9	Row...	Ruby	5
10	Row...	Groovy	5



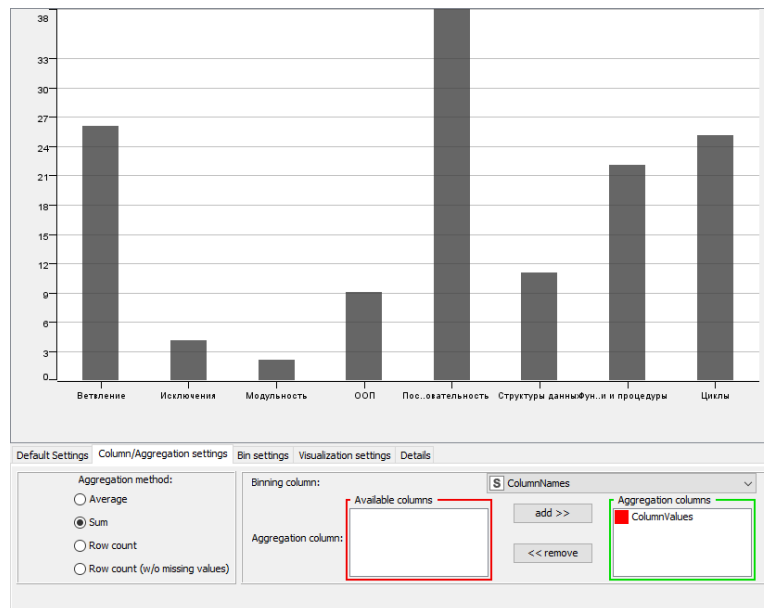
3. Знание концепций программирования

За обработку данных результатов отвечает следующий фрагмент схемы (выбирает и обрабатывает необходимые данные из исходной таблицы, а затем строит диаграмму):



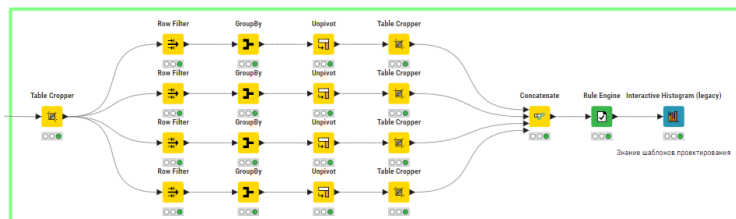
На выходе получаем таблицу со значениями (количеством выборов для каждого варианта) и диаграмму распределения:

#	Row...	ColumnNames String	ColumnValues Number (integer)
1	Row0	Последовательность	38
2	Row...	Ветвление	26
3	Row...	Циклы	25
4	Row...	Функции и процедуры	22
5	Row...	Структуры данных	11
6	Row...	ООП	9
7	Row...	Исключения	4
8	Row...	Модульность	2



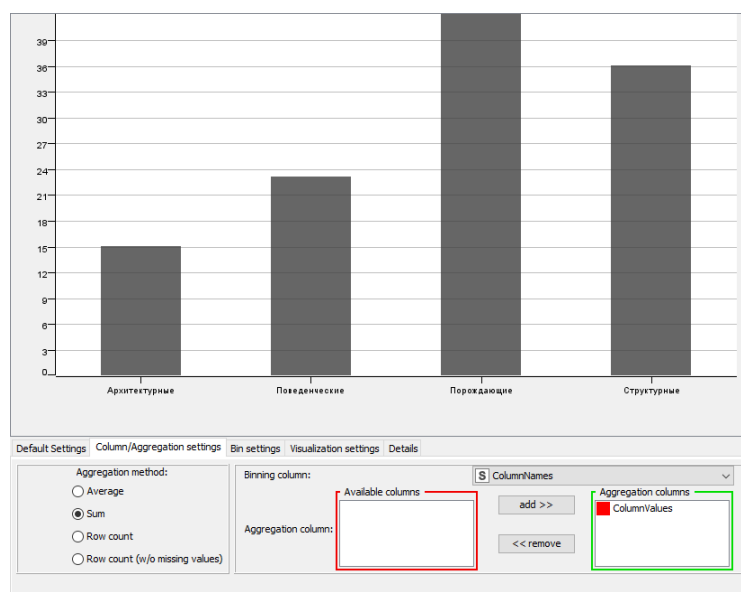
4. Знание шаблонов проектирования

За обработку данных результатов отвечает следующий фрагмент схемы (выбирает и обрабатывает необходимые данные из исходной таблицы, а затем строит диаграмму):



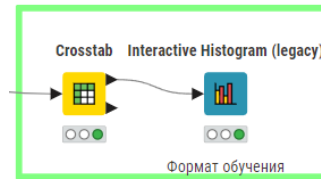
На выходе получаем таблицу со значениями (количеством выборов для каждого варианта) и диаграмму распределения:

#	Row...	ColumnNames String	ColumnValues Number (integer)
1	Row0	Порождающие	42
2	Row...	Структурные	36
3	Row...	Поведенческие	23
4	Row...	Архитектурные	15

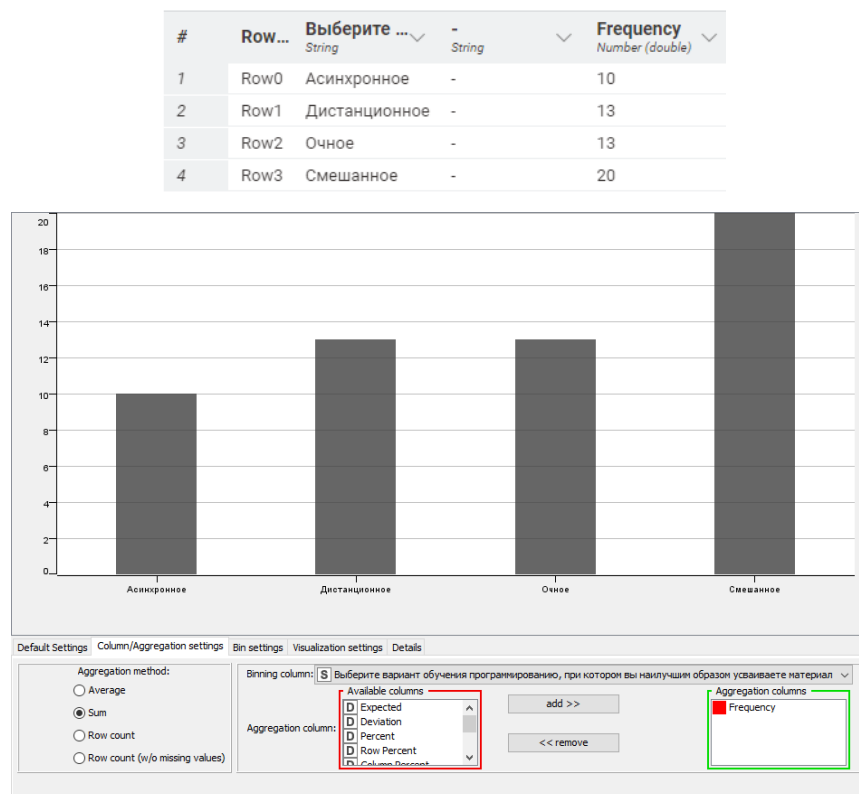


5. Предпочитаемый формат обучения

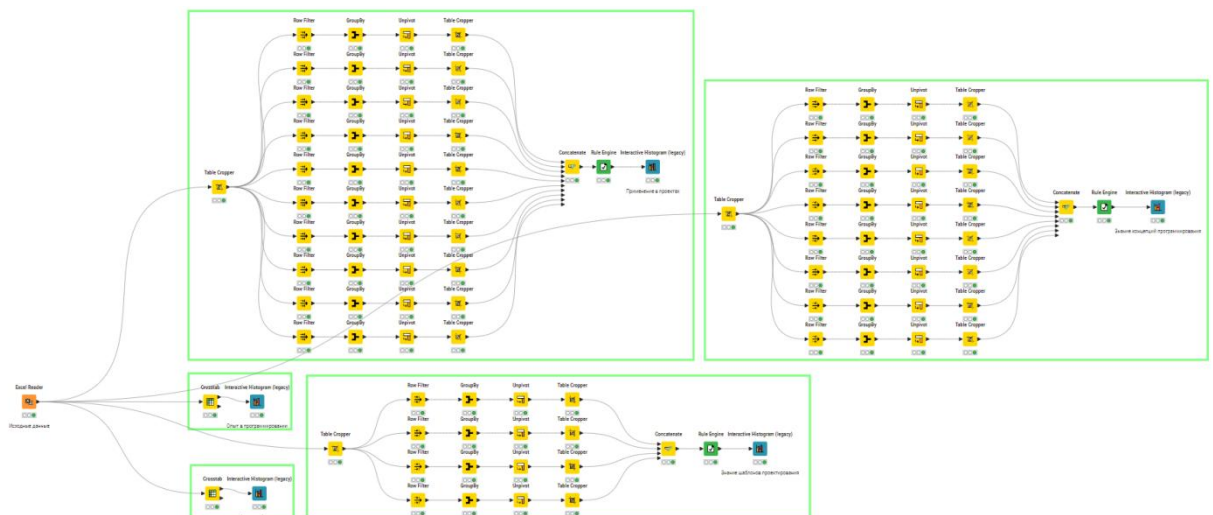
За обработку данных результатов отвечает следующий фрагмент схемы (выбирает необходимые данные из исходной таблицы, а затем строит диаграмму):



На выходе получаем таблицу со значениями (количеством выборов для каждого варианта) и диаграмму распределения:



Общая схема обработки для первых пяти пунктов



Выводы по первым пяти вопросам:

1. Большая часть студентов, на момент прохождения опроса, занимается программированием на протяжении 3 или 4 лет.
2. Для написания проектов и реализации сложных задач подавляющее большинство студентов применяет язык программирования Python.
3. Наиболее известными концепциями программирования для студентов являются: последовательность, ветвление, циклы, а также функции и процедуры.
4. Наиболее известными шаблонами проектирования для студентов являются порождающие паттерны и структурные паттерны.
5. Наиболее предпочтительным форматом обучения для студентов является смешанный вариант обучения.

Теперь рассмотрим обработку результатов для получения ответов на 6 и 7 вопросы, в которых необходимо определить коррелируют ли определенные распределения.

Обработка через Excel

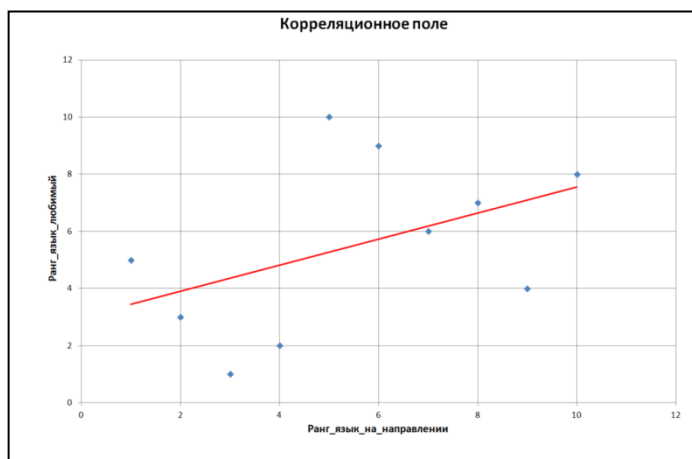
(Файл со всеми таблицами и диаграммами приложен в ответе)

6. Корреляция между основными языками программирования на направлениях и любимыми языками программирования у студентов.

Таблица распределения голосов по вариантам и ранги вариантов (Лист “Корреляция ”):

	A	B	C	D	E	F	G
1	Язык	Количество_направление	Количество_любимый	Ранг_направление	Ранг_любимый	d	d^2
2	Python	9	13	3	1	2	4
3	JavaScript	8	9	4	2	2	4
4	Java	3	5	7	6	1	1
5	C#	2	4	8	7	1	1
6	C++	11	8	2	3	-1	1
7	C	0	3	10	8	2	4
8	Swift	1	7	9	4	5	25
9	Kotlin	13	6	1	5	-4	16
10	Ruby	4	1	6	9	-3	9
11	Groovy	5	0	5	10	-5	25
12						Сумма:	90

По столбцам, содержащим ранги, было построено корреляционное поле (Лист “Корреляционное_поле_1 ”):



После получения из таблицы ответов, данной таблицы, был вычислен ранг корреляции двумя способами, а затем с использованием ранга корреляции был найден Т-критерий, данные значения были занесены в таблицу (Лист “Корреляция ”):

Количество вводимых параметров:	10
Ранг корреляции	
Способ 1 (постепенное вычисление):	0,455
Способ 2 (функция КОРРЕЛ):	0,455
t	1,443

По таблице критических значений было найдено, что при количестве степеней свободы $n - 2 = 8$ и уровне значимости 0,05, $t_{кр} = 2,31$.

Значит $t < t_{кр}$, а значит нулевая гипотеза не отвергается.

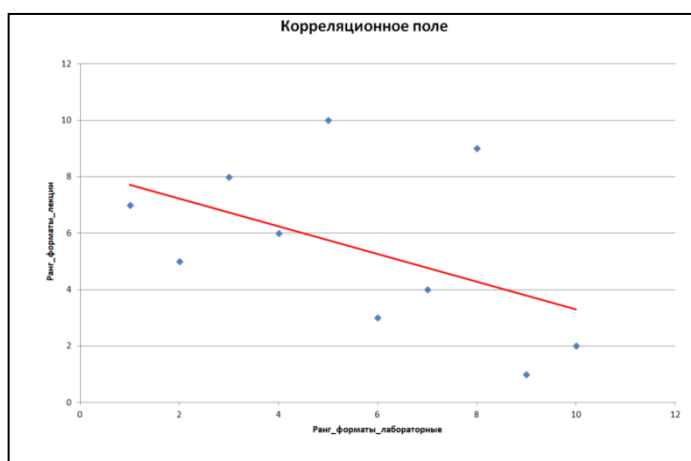
Таким образом, корреляция между языками программирования, выбранными на направлении обучения и любимыми языками программирования студентов, отсутствует или близка к нулю.

7. Корреляция между выбранными форматами представления информации по лекционным занятиям и лабораторным.

Таблица распределения голосов по вариантам и ранги вариантов (Лист “Корреляция ”):

Форма	Количество_лекции	Количество_лабораторные	Ранг_лекции	Ранг_лабораторные	d	d^2
Текстовы	3	6	7	4	3	9
Презента	5	0	5	10	-5	25
Видеозап	15	3	1	7	-6	36
Скринкас	12	5	2	5	-3	9
Программ	0	11	10	2	8	64
Полноцен	6	4	4	6	-2	4
Онлайн р	8	2	3	8	-5	25
Борд в Re	1	15	9	1	8	64
Страница	4	9	6	3	3	9
Курс на о	2	1	8	9	-1	1
Сумма:					246	

По столбцам, содержащим ранги, было построено корреляционное поле (Лист “Корреляционное_поле_2 ”):



После получения из таблицы ответов, данной таблицы, был вычислен ранг корреляции двумя способами, а затем с использованием ранга корреляции был найден Т-критерий, данные значения были занесены в таблицу (Лист “Корреляция ”):

Количество вводимых параметров:	10
Ранг корреляции	
Способ 1 (постепенное вычисление):	-0,491
Способ 2 (функция КОРРЕЛ):	-0,491
t	1,594

По таблице критических значений было найдено, что при количестве степеней свободы $n - 2 = 8$ и уровне значимости 0,05, $t_{кр} = 2,31$.

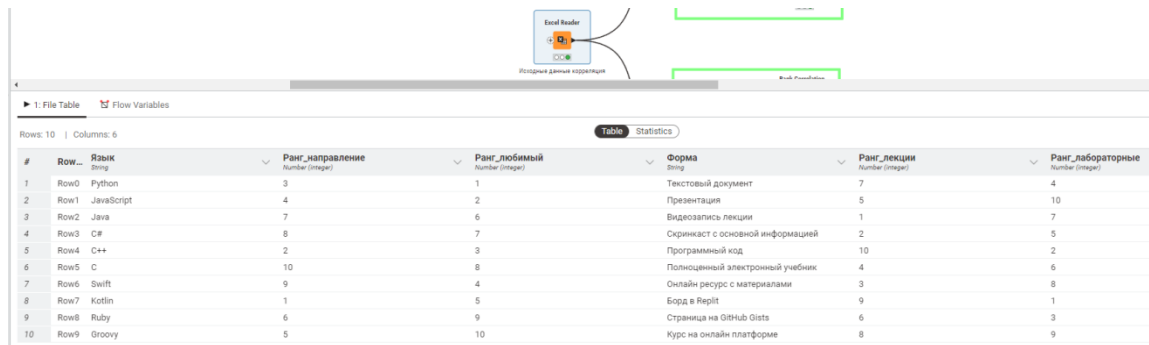
Значит $t < t_{кр}$, а значит нулевая гипотеза не отвергается.

Таким образом, корреляция между выбранными форматами представления информации по лекционным занятиям и лабораторным, отсутствует или близка к нулю.

Обработка через KNIME

(Архив с разработанным с помощью KNIME проектом приложен в ответе)

С помощью “Excel Reader” в проект были загружены данные из сводной таблицы с рангами для проведения двух корреляционных анализов:

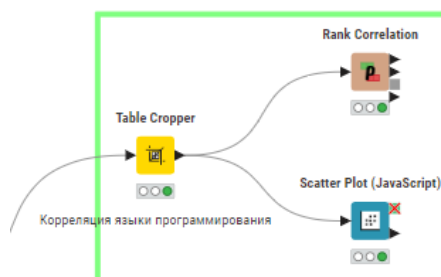


#	Row	Язык String	Ранг_направление Number (Integer)	Ранг_любимый Number (Integer)	Форма String	Ранг_лекции Number (Integer)	Ранг_лабораторные Number (Integer)
1	Row0	Python	3	1	Текстовый документ	7	4
2	Row1	JavaScript	4	2	Презентация	5	10
3	Row2	Java	7	6	Видеозапись лекции	1	7
4	Row3	C#	8	7	Скринкаст с основной информацией	2	5
5	Row4	C++	2	3	Программный код	10	2
6	Row5	C	10	8	Полноценный электронный учебник	4	6
7	Row6	Swift	9	4	Онлайн ресурс с материалами	3	8
8	Row7	Kotlin	1	5	Борд в Replit	9	1
9	Row8	Ruby	6	9	Страница на GitHub Gists	6	3
10	Row9	Groovy	5	10	Курс на онлайн платформе	8	9

Затем была проведена обработка результатов для каждого вопроса и составлена общая схема.

6. Корреляция между основными языками программирования на направлениях и любимыми языками программирования у студентов.

За обработку данных результатов отвечает следующий фрагмент схемы (выбирает и обрабатывает необходимые данные из исходной таблицы, а затем строит корреляционное поле и находит ранг корреляции):

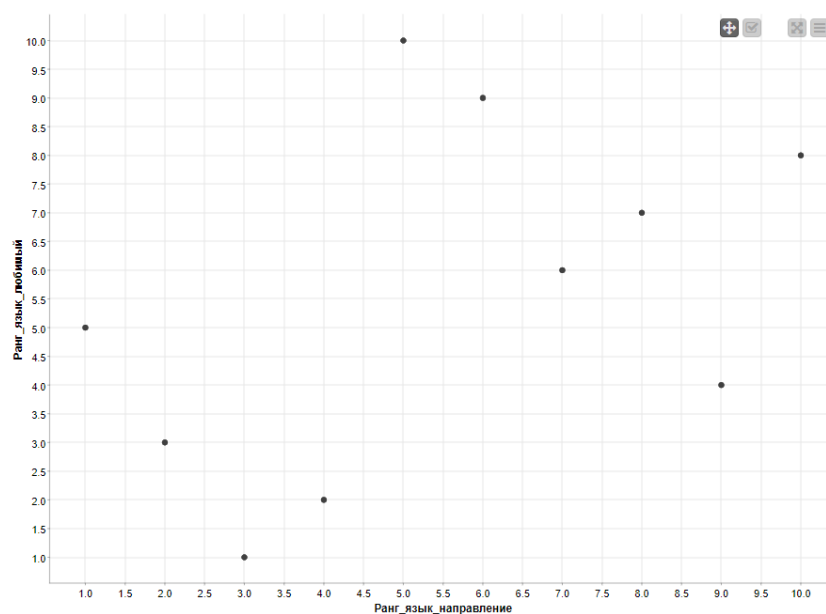


Данные выбранные для построения корреляционного поля и нахождения ранга корреляции в данном пункте:

#	Row...	Язык String	Ранг_направление Number (integer)	Ранг_любимый Number (integer)
1	Row0	Python	3	1
2	Row1	JavaScript	4	2
3	Row2	Java	7	6
4	Row3	C#	8	7
5	Row4	C++	2	3
6	Row5	C	10	8
7	Row6	Swift	9	4
8	Row7	Kotlin	1	5
9	Row8	Ruby	6	9
10	Row9	Groovy	5	10

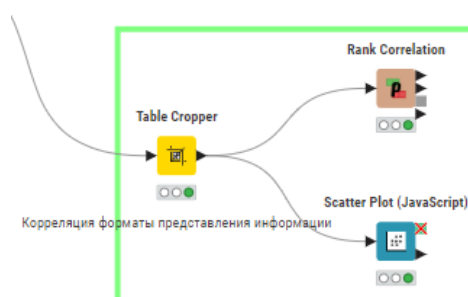
На выходе получаем таблицу со значением ранга корреляции и корреляционное поле:

#	Row...	First column name String	Second column name String	Correlation value Number (double)	p value Number (double)	Degrees of freedom Number (integer)
1	Row0	Ранг_направление	Ранг_любимый	0.455	0.187	8



7. Корреляция между wybranymi форматами представления информации по лекционным занятиям и лабораторным.

За обработку данных результатов отвечает следующий фрагмент схемы (выбирает и обрабатывает необходимые данные из исходной таблицы, а затем строит корреляционное поле и находит ранг корреляции):

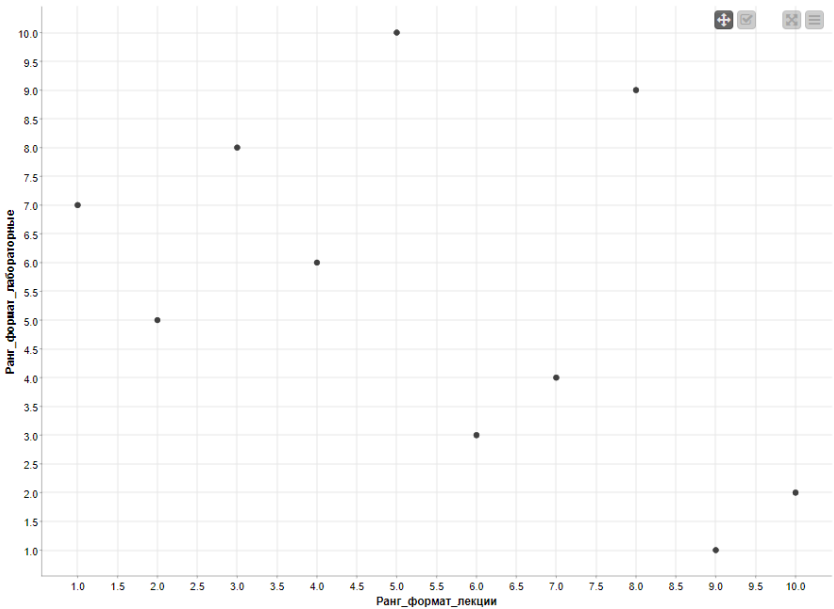


Данные выбранные для построения корреляционного поля и нахождения ранга корреляции в данном пункте:

#	Row...	Форма <small>String</small>	Ранг_лекции <small>Number (integer)</small>	Ранг_лабораторные <small>Number (integer)</small>
1	Row0	Текстовый документ	7	4
2	Row1	Презентация	5	10
3	Row2	Видеозапись лекции	1	7
4	Row3	Скринкаст с основной информацией	2	5
5	Row4	Программный код	10	2
6	Row5	Полноценный электронный учебник	4	6
7	Row6	Онлайн ресурс с материалами	3	8
8	Row7	Борд в Replit	9	1
9	Row8	Страница на GitHub Gists	6	3
10	Row9	Курс на онлайн платформе	8	9

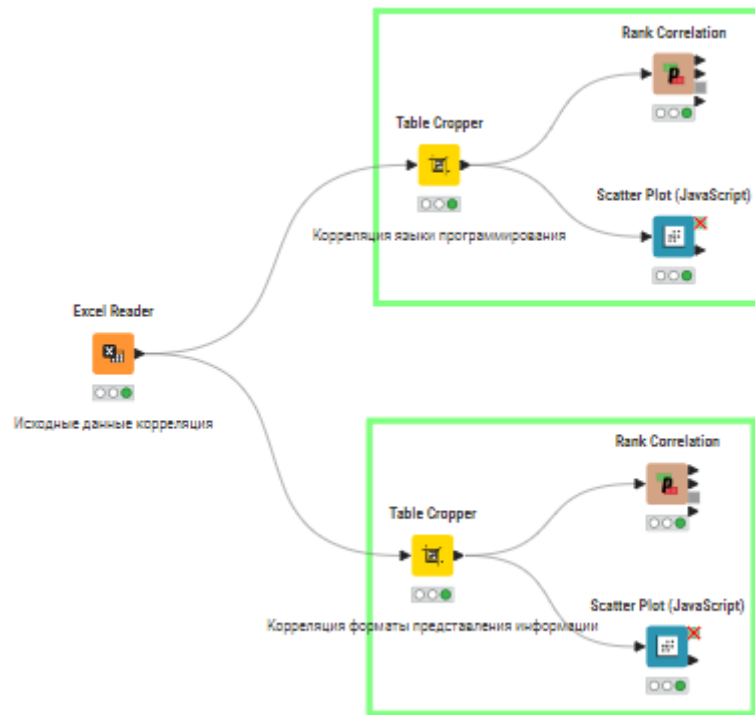
На выходе получаем таблицу со значением ранга корреляции и корреляционное поле:

#	Row...	First column name <small>String</small>	Second column name <small>String</small>	Correlation value <small>Number (double)</small>	p value <small>Number (double)</small>	Degrees of freedom <small>Number (integer)</small>
1	Row0	Ранг_лекции	Ранг_лабораторные	-0.491	0.15	8



Проведя аналогичные рассуждения для результатов, полученных в пунктах 6 и 7 с помощью KNIME, мы придём к тем же выводам, что и в пунктах 6 и 7 при обработке с помощью Excel.

Общая схема обработки для шестого и седьмого пунктов



Вывод

В данной лабораторной работе мной была реализована обработка результатов проведенного опроса, по нескольким важным пунктам. В ходе обработки результатов были построены распределения значений в виде сводных таблиц и диаграмм, а также проведен корреляционный анализ для двух пар распределений. Данная работа была выполнена в двух программах, позволяющих провести обработку данных. А именно в EXCEL и KNIME. В ходе выполнения данной лабораторной работы, я научился загружать данные в KNIME, а затем, с помощью его встроенных средств, научился обрабатывать загруженные данные и строить по обработанным данным различные диаграммы.