



Введение в анализ данных

Лекция 6.2

Шевцов Василий Викторович,
директор ДИТ РУДН, shevtsov_vv@rudn.university

Визуализация данных

Проектная диаграмма Ганта. Диаграмма сравнений Торнадо. Каскадная диаграмма (диаграмма отклонений Водопад).

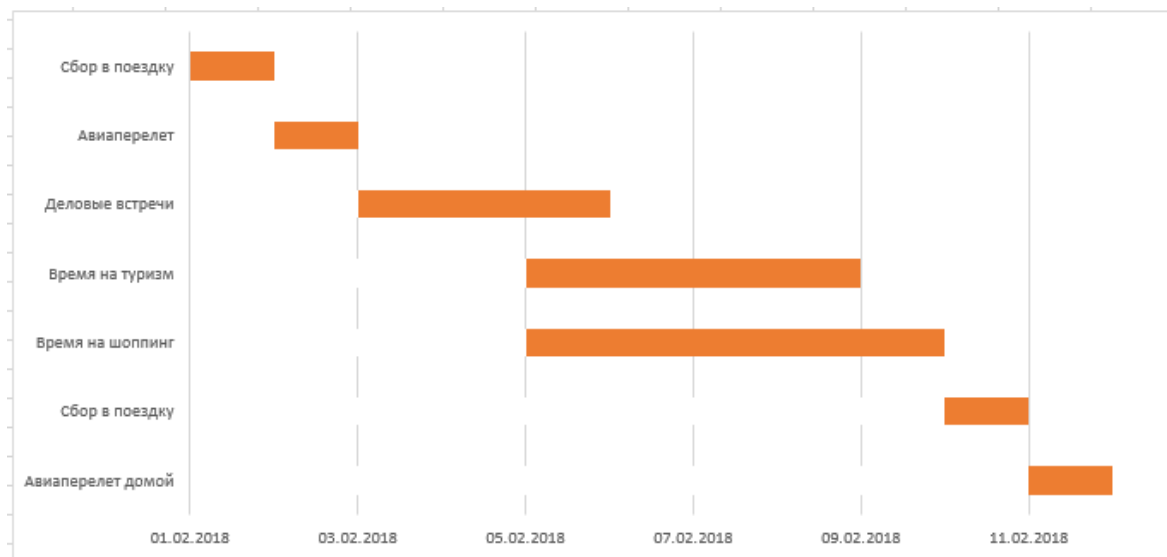
Проектная диаграмма Ганта

Проектная диаграмма Ганта

Диаграмма Ганта названа в честь Генри Гантта, американского инженера-механика и консультанта по вопросам управления, который изобрел этот график еще в 1910-х годах.

Диаграмма Ганта в Excel представляет собой проекты или задачи в виде каскадных горизонтальных гистограмм.

Диаграмма Ганта иллюстрирует структуру разбивки проекта, отображая даты начала и окончания, а также различные отношения между действиями проекта и таким образом помогает отслеживать задачи по сравнению с запланированным временем или predetermined вехами.



Исходные данные

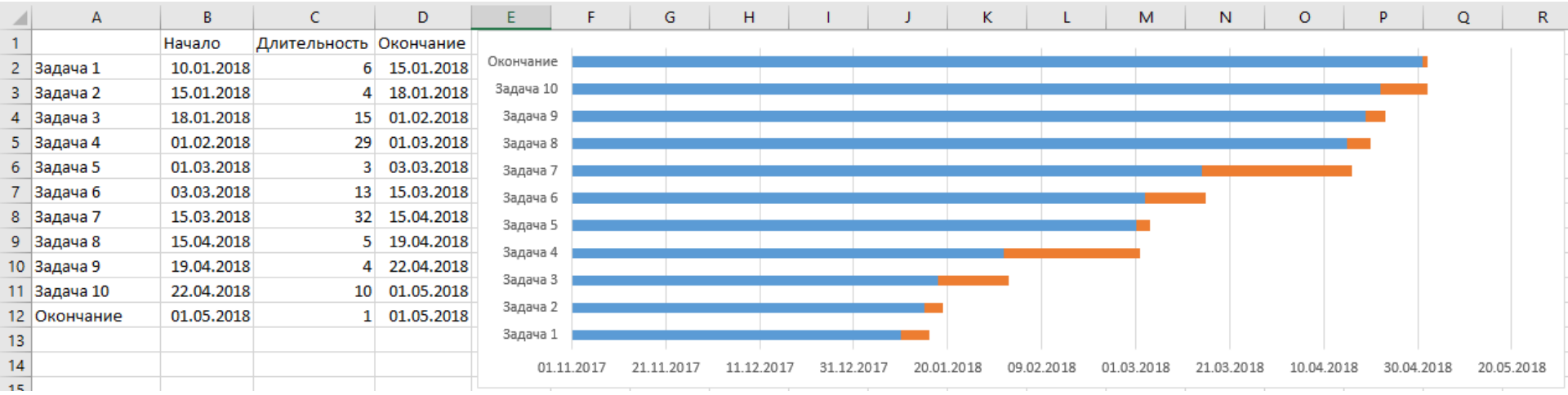
	A	B	C	D
1		Начало	Длительность	Окончание
2	Задача 1	10.01.2018	6	15.01.2018
3	Задача 2	15.01.2018	4	18.01.2018
4	Задача 3	18.01.2018	15	01.02.2018
5	Задача 4	01.02.2018	29	01.03.2018
6	Задача 5	01.03.2018	3	03.03.2018
7	Задача 6	03.03.2018	13	15.03.2018
8	Задача 7	15.03.2018	32	15.04.2018
9	Задача 8	15.04.2018	5	19.04.2018
10	Задача 9	19.04.2018	4	22.04.2018
11	Задача 10	22.04.2018	10	01.05.2018
12	Окончание	01.05.2018	1	01.05.2018

Необходимы два поля:

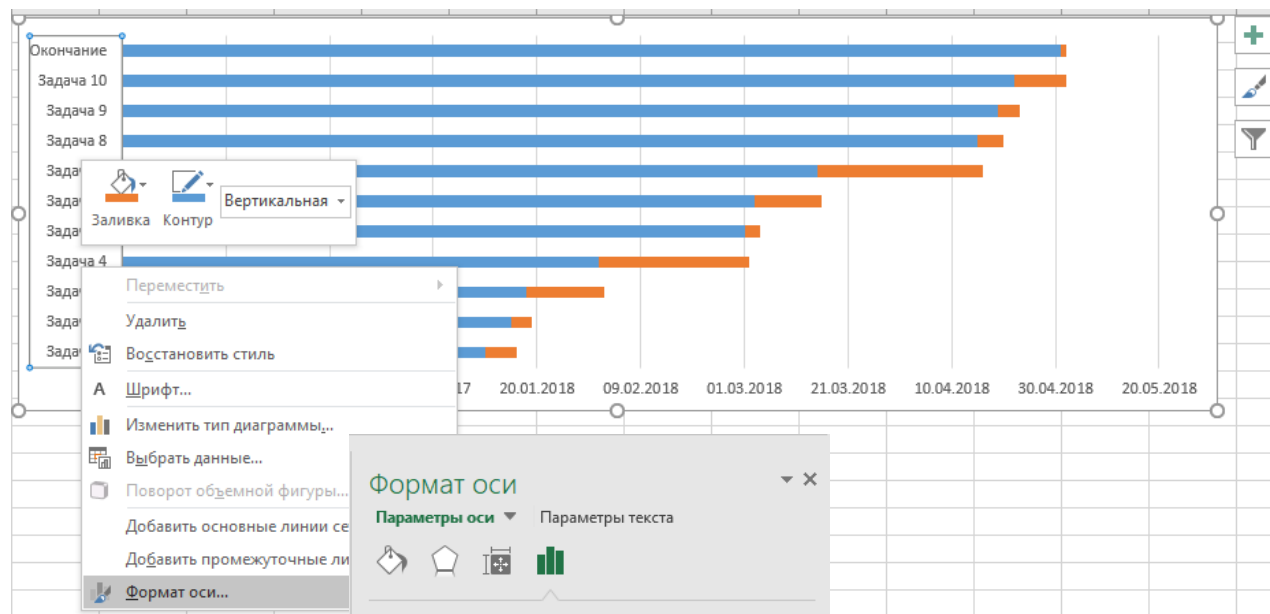
- Дата (время) начала события
- Длительность события

Начало события будет определять положение периода длительности на временной шкале

Строится линейчатая диаграмма с накоплением.
В данном примере синяя область – дата начала, желтая - длительность



Проектная диаграмма Ганта. Окончание настройки



- Обратный порядок категорий
- Минимальны и максимальные значения
- Параметры оси.

Формат оси

Параметры оси | Параметры текста

Параметры оси

Тип оси

- ☒ Автовыбор на основе данных
- ☐ Ось текста
- ☐ Ось дат

Горизонтальная ось пересекает

- ☒ Автовыбор
- ☐ в номере категории
- ☐ в максимальной категории

Положение оси

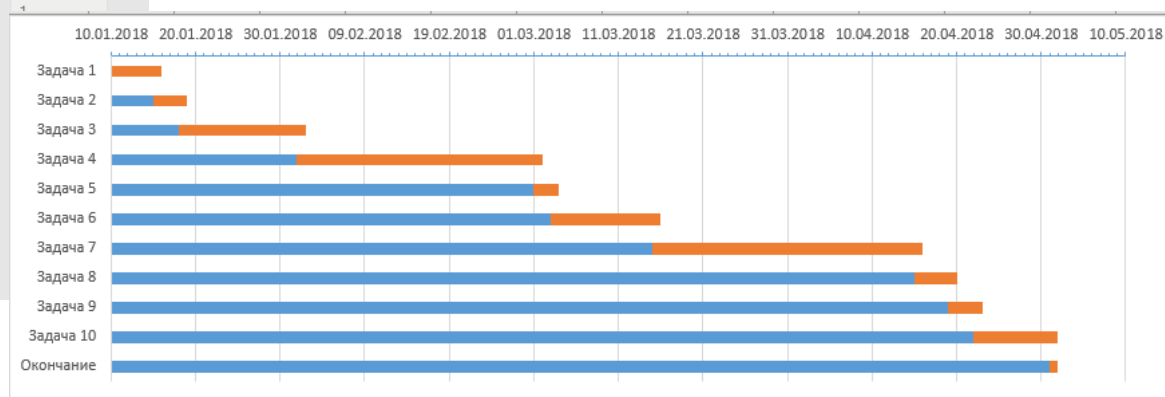
- ☐ по делениям
- ☒ между делениями

☒ Обратный порядок категорий

Деления

Подписи

Число



Проектная диаграмма Ганта. Окончание настройки

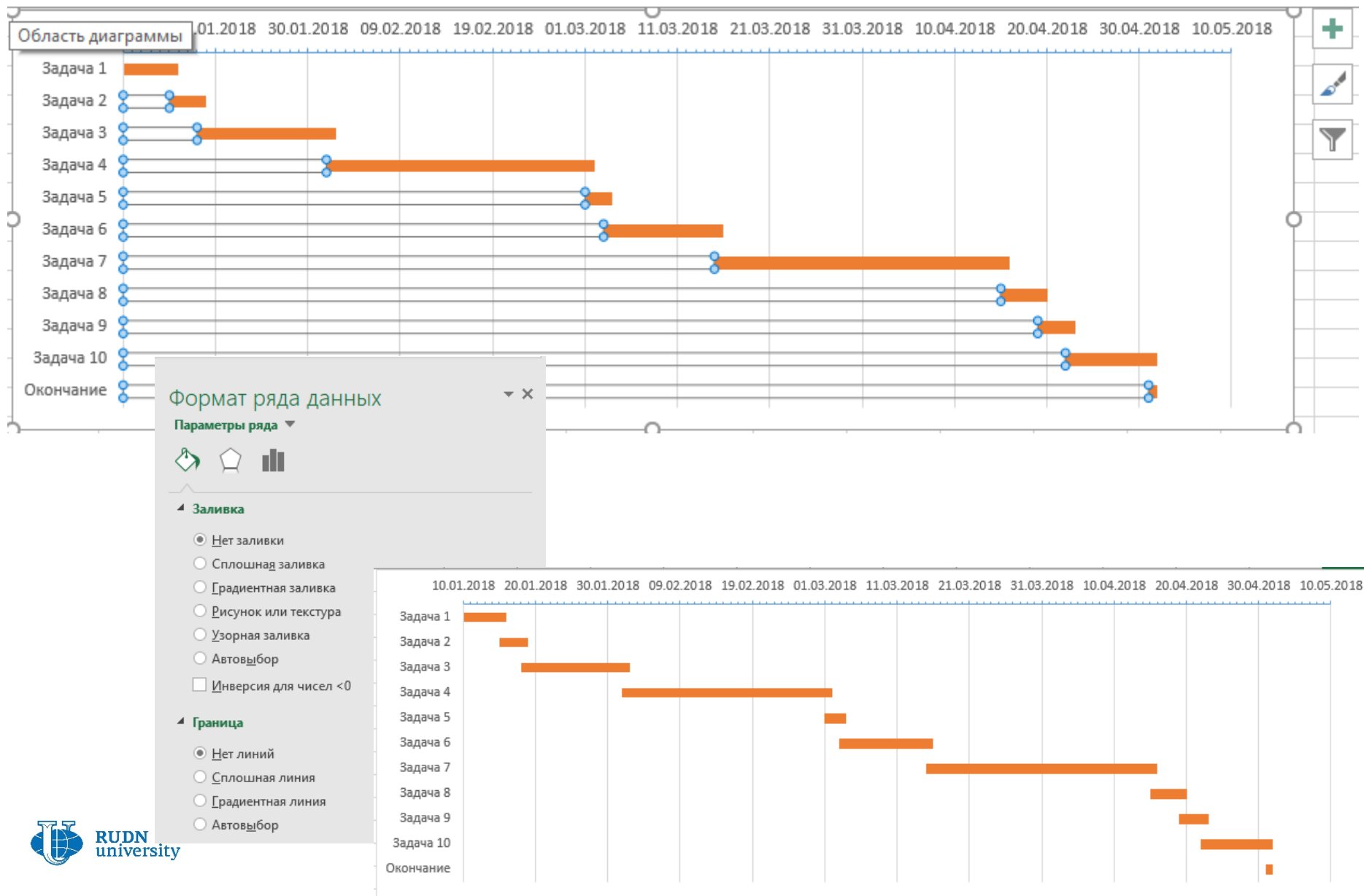


Диаграмма сравнений Торнадо

Диаграмма сравнений Торнадо

В зарубежных источниках такой тип диаграммы называют tornado-chart. От наших соотечественников я слышал варианты "диаграмма-горшок" и "диаграмма-бабочка" - у кого на что больше похоже :)

Такой тип диаграмм удобно использовать для наглядного сравнение двух наборов данных, например:

мы и конкуренты по основным направлениям (товарам, услугам)
количество мужчин и женщин разного возраста среди покупателей или зрителей
этот год и прошлый год по основным KPI и т.д.

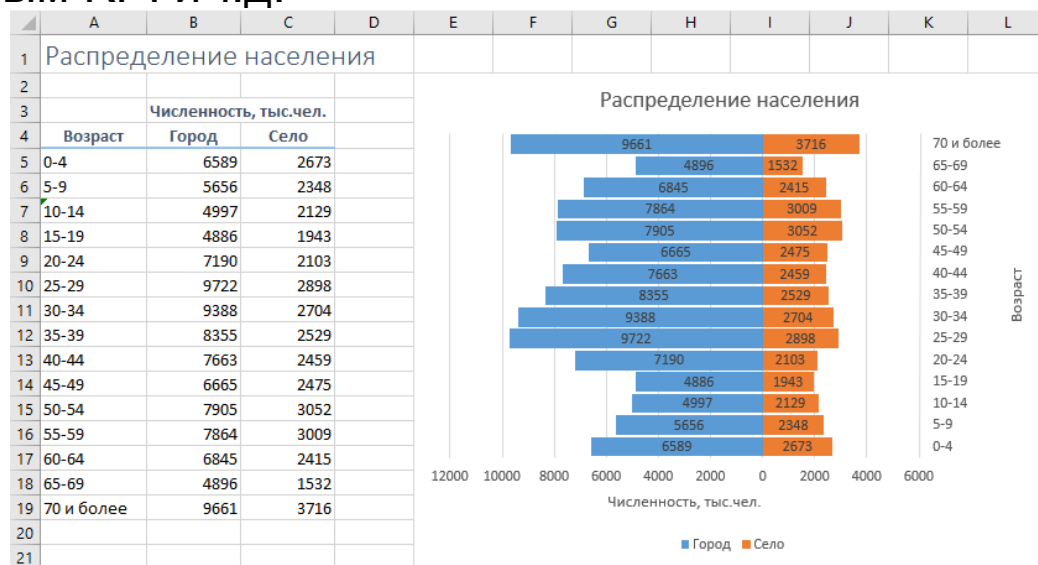


Диаграмма сравнений Торнадо

	A	B	C
1	Факультет	Мужской	Женский
2	Аграрно-технологический институт	423	855
3	Гуманитарных и социальных наук	754	1049
4	Инженерная академия	1510	767
5	Институт биохимической технологии и нанотехнологии	15	14
6	Институт гостиничного бизнеса и туризма	168	383
7	Институт иностранных языков	182	936
8	Институт космических технологий	21	8
9	Институт мировой экономики и бизнеса	223	340
10	Кафедра сравнительной образовательной политики	3	5
11	Медицинский институт	1834	3237
12	Учебно-научный институт гравитации и космологии	10	4

Чтобы оппозитно расположить значения, необходимо одну категорию домножить на -1. Это можно сделать с помощью вспомогательного столбца (-1) и специальной вставки.

	A	B	C
1	Факультет	Мужской	Женский
2	Аграрно-технологический институт	-423	855
3	Гуманитарных и социальных наук	-754	1049
4	Инженерная академия	-1510	767
5	Институт биохимической технологии и нанотехнологии	-15	14
6	Институт гостиничного бизнеса и туризма	-168	383
7	Институт иностранных языков	-182	936
8	Институт космических технологий	-21	8
9	Институт мировой экономики и бизнеса	-223	340
10	Кафедра сравнительной образовательной политики	-3	5
11	Медицинский институт	-1834	3237
12	Учебно-научный институт гравитации и космологии	-10	4
13	Физико-математических и естественных наук	-586	480
14	Филологический	-541	1914
15	Экологический	-146	254
16	Экономический	-988	1278
17	Юридический институт	-719	768

Специальная вставка

Вставить

☒ все
 ☐ с исходной темой

☐ формулы
 ☐ без рамки

☐ значения
 ☐ ширины столбцов

☐ форматы
 ☐ формулы и форматы чисел

☐ примечания
 ☐ значения и форматы чисел

☐ условия на значения
 ☐ все условные форматы объединить

Операция

☐ нет
 ☒ умножить

☐ сложить
 ☐ разделить

☐ вычесть
 ☐

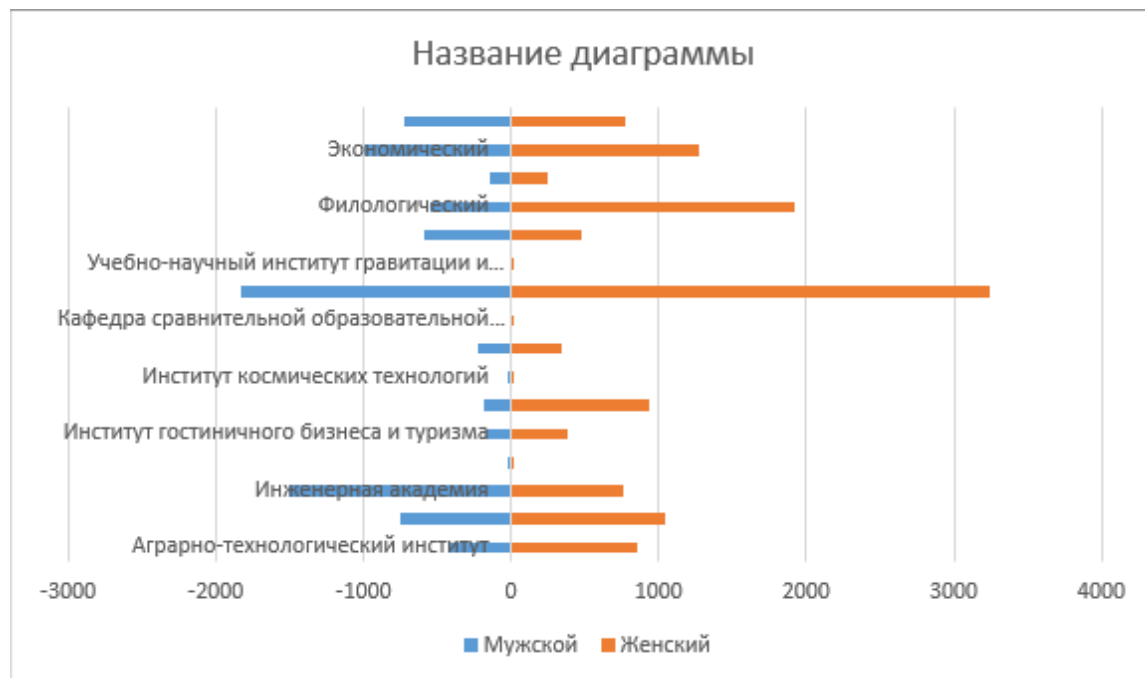
☐ пропускать пустые ячейки
 ☐ транспонировать

Вставить связь

OK

Отмена

Диаграмма сравнений Торнадо



Недостатки:

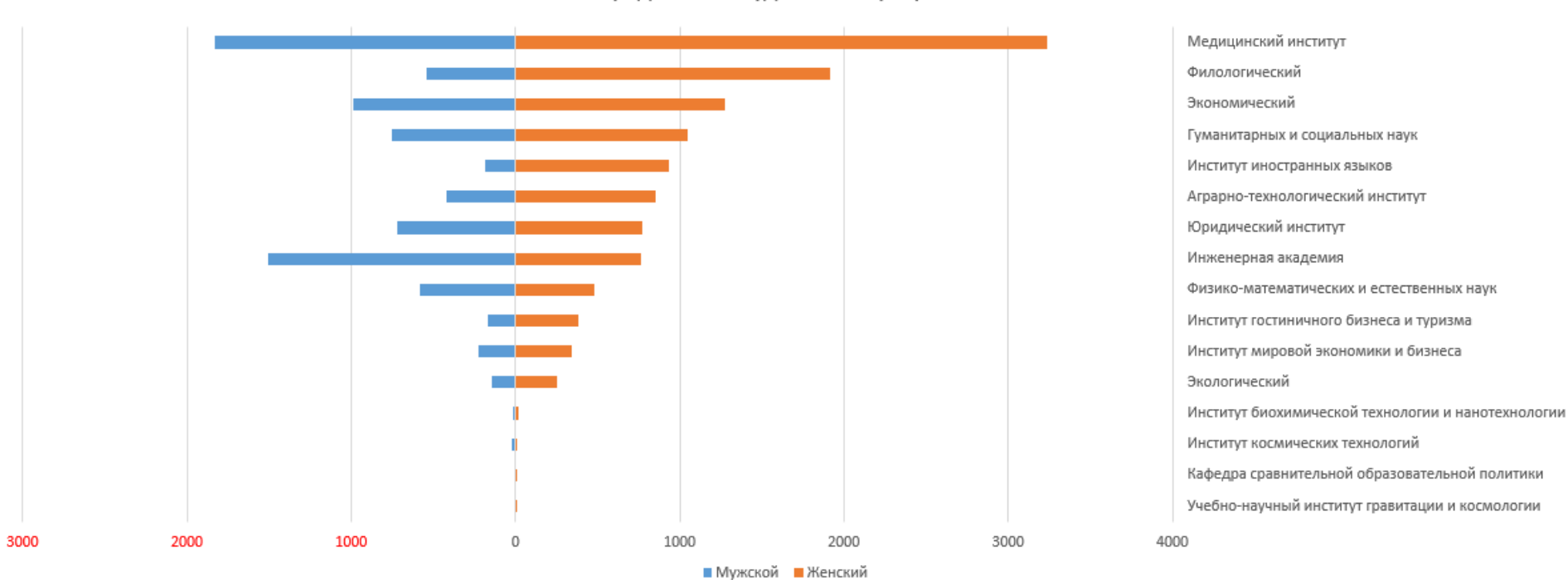
- В одной из категорий отрицательные значения.
- Мешают подписи оси.

Решение:

- Выбирается числовой формат, который для отрицательных значений не выводит знак минус. Это или готовый формат, или пользовательский.
- В свойствах оси выбирается расположение подписей.

Диаграмма сравнений Торнадо. Итог

Распределение студентов на факультетах

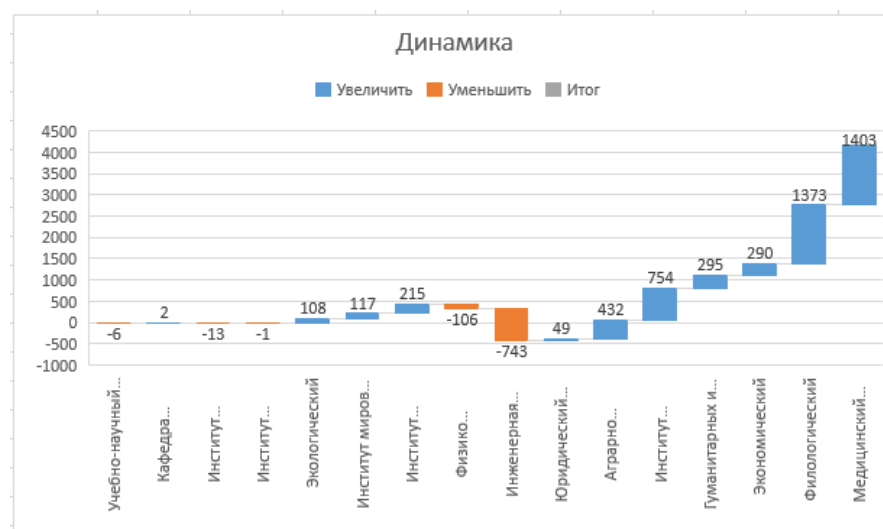
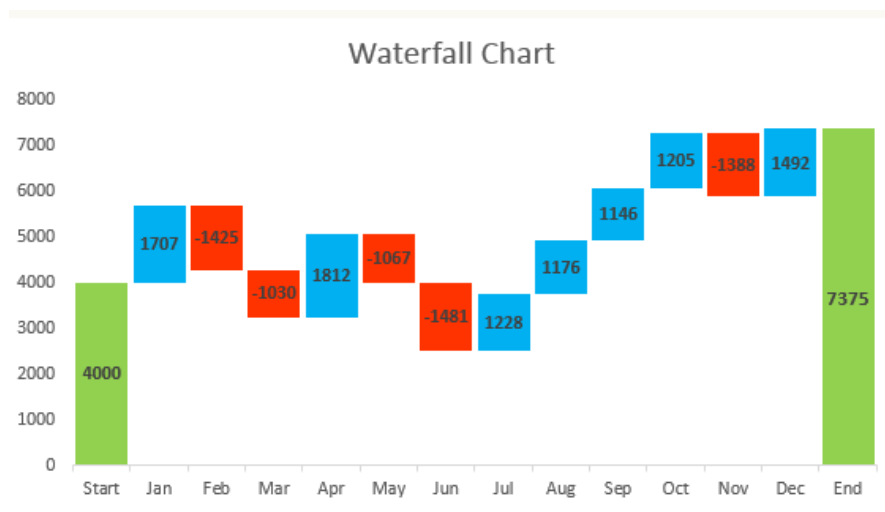


- Выбран готовый числовой формат.
- Значения оси отсортированы по убыванию.
- Изменено расположение подписей.

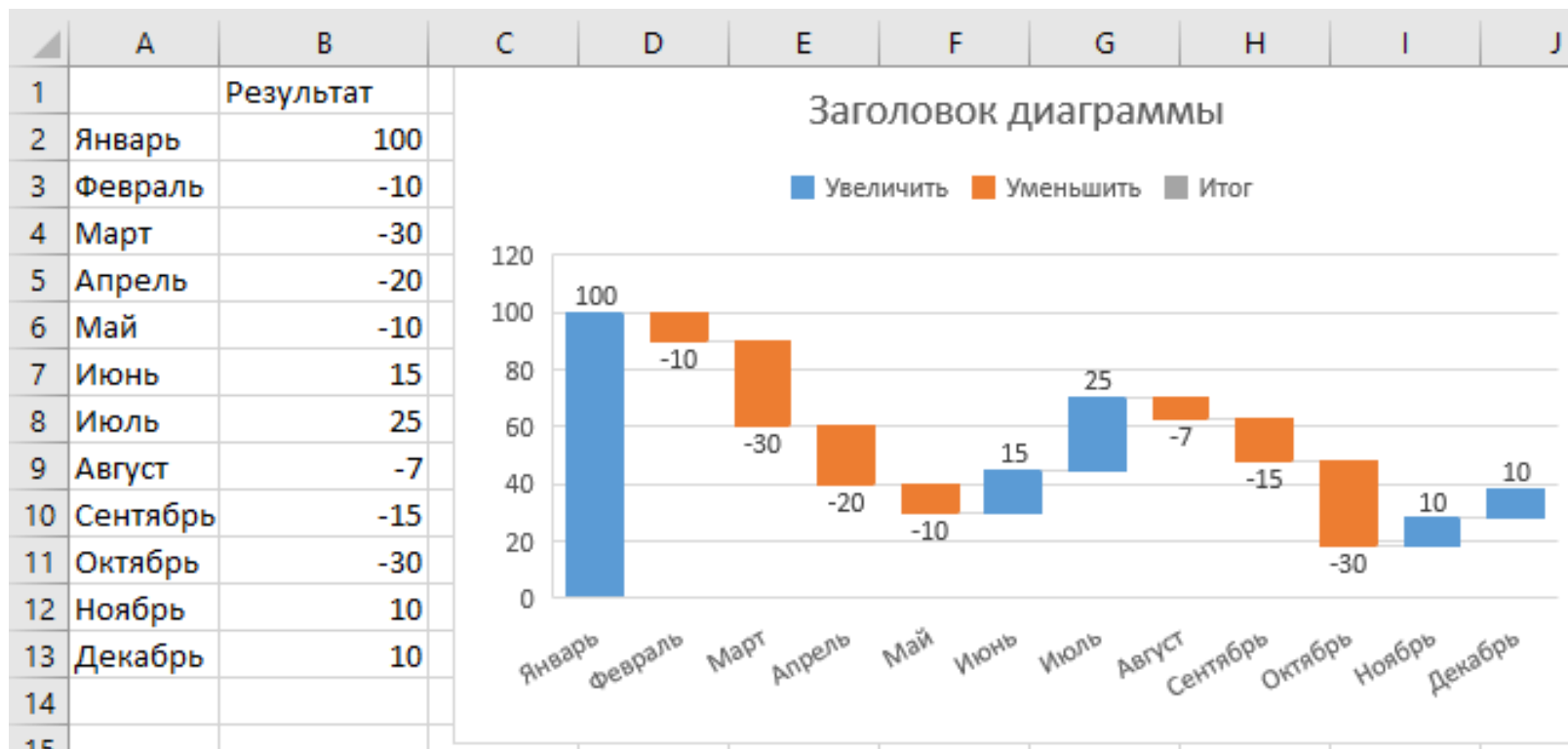
Каскадная диаграмма

Каскадная диаграмма

Диаграмма "Водопад" – это особый тип диаграммы в Excel. Обычно используется для того, чтобы показать, как исходные данные увеличиваются или уменьшаются в результате ряда изменений. Первый и последний столбцы в обычной каскадной диаграмме показывают суммарные значения. Промежуточные колонки являются плавающими и обычно показывают положительные или отрицательные изменения, происходящие от одного периода к другому, а в конечном итоге общий результат, т.е. суммарное значение. Как правило, колонки окрашены в разные цвета, чтобы наглядно выделить положительные и отрицательные значения.



Каскадная диаграмма в Excel 2016



В Microsoft Office Excel 2016 каскадная диаграмма добавлена в набор диаграмм.

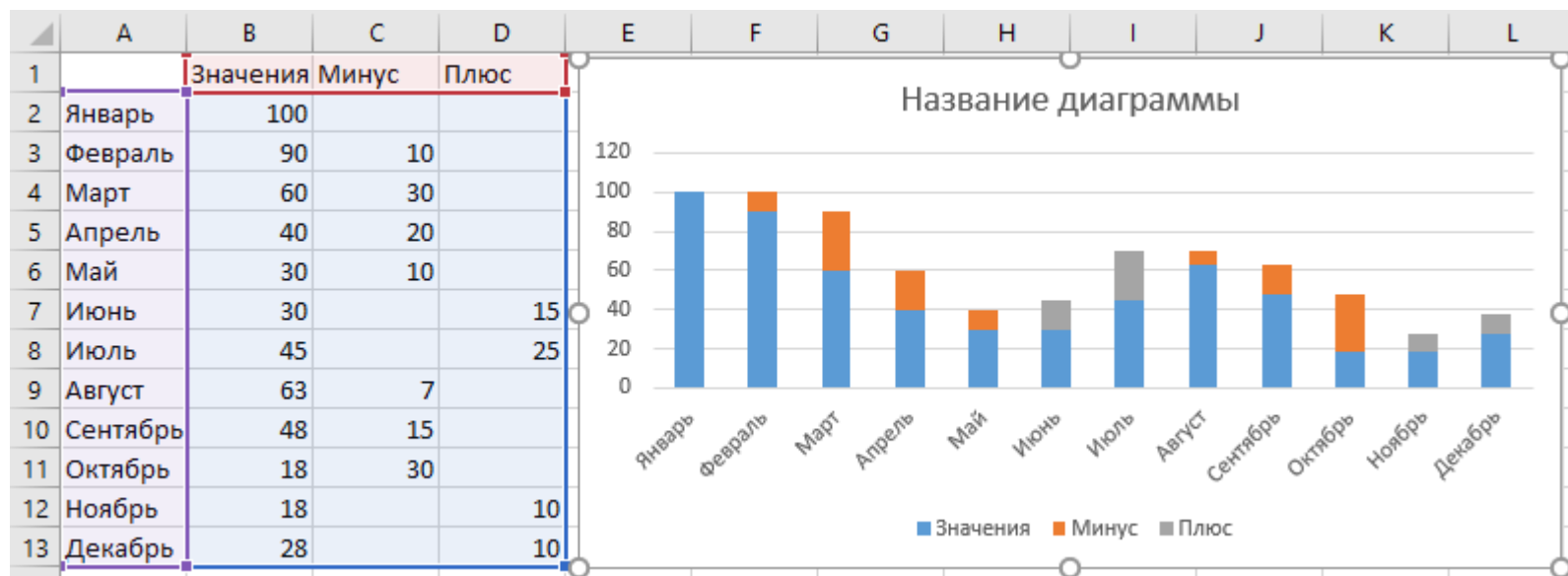
В более ранних версиях данный вид диаграммы отсутствует.

Каскадная диаграмма в Excel 2013

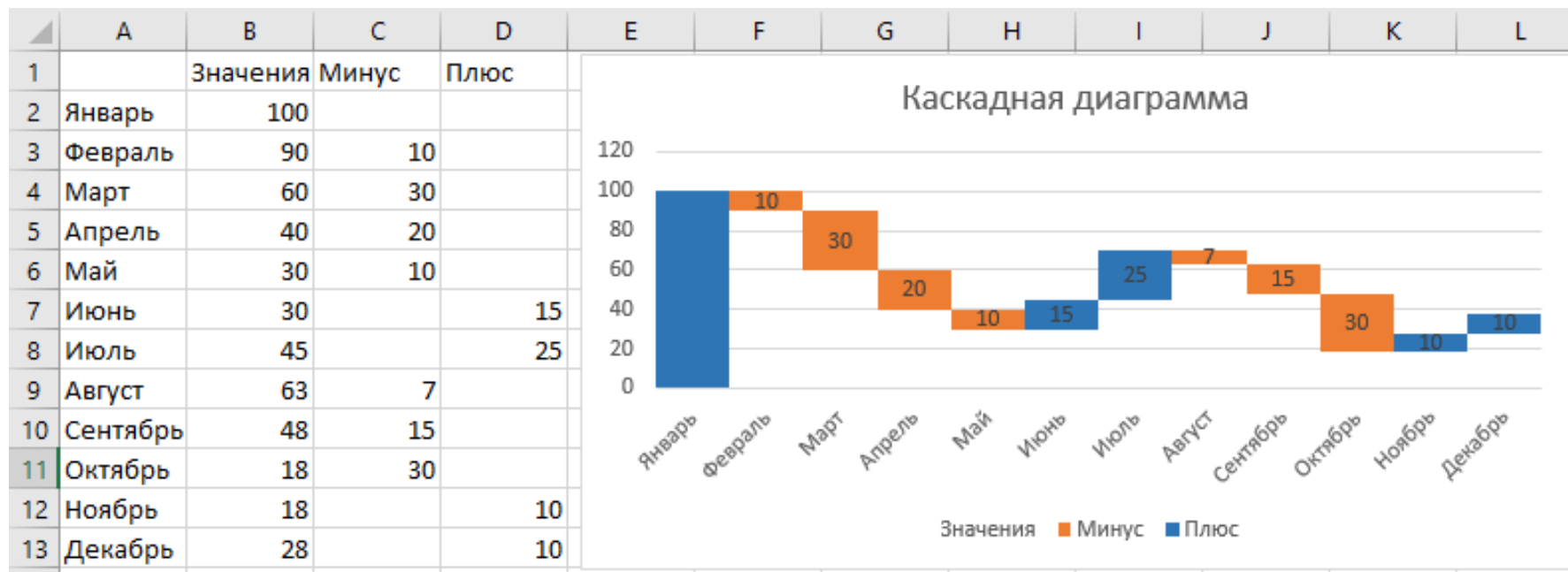
	A	B	C	D
1		Значения	Минус	Плюс
2	Январь	100		
3	Февраль	90	10	
4	Март	60	30	
5	Апрель	40	20	
6	Май	30	10	
7	Июнь	30		15
8	Июль	45		25
9	Август	63	7	
10	Сентябрь	48	15	
11	Октябрь	18	30	
12	Ноябрь	18		10
13	Декабрь	28		10

В исходных данных нужны абсолютные значения и отклонения, размещенные в отдельных столбцах для образования отдельных рядов данных

Создается столбчатая диаграмма с накоплением



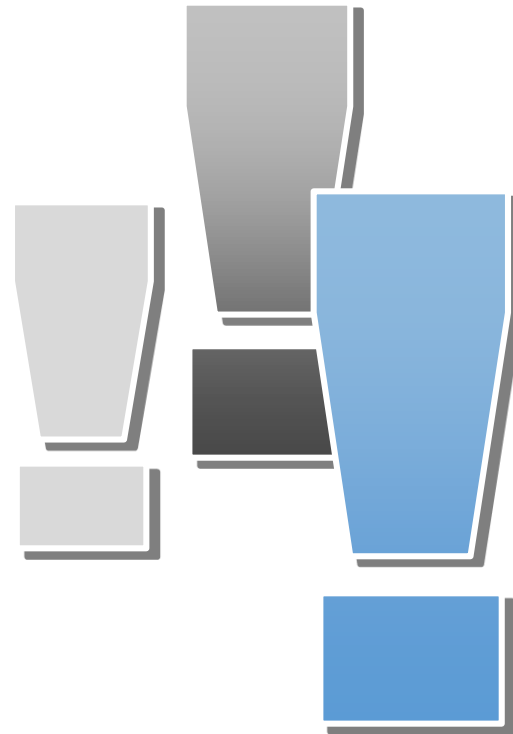
Каскадная диаграмма в Excel 2013



Окончание настройки:

- Ряд **Значения** надо сделать невидимым
- Первое значение в ряде **Значения** залить цветом положительных значений
- Уменьшить или убрать боковой зазор между столбцами
- Добавить подписи данных

Спасибо за внимание!



Шевцов Василий Викторович

shevtsov_vv@rudn.university
+7(903)144-53-57