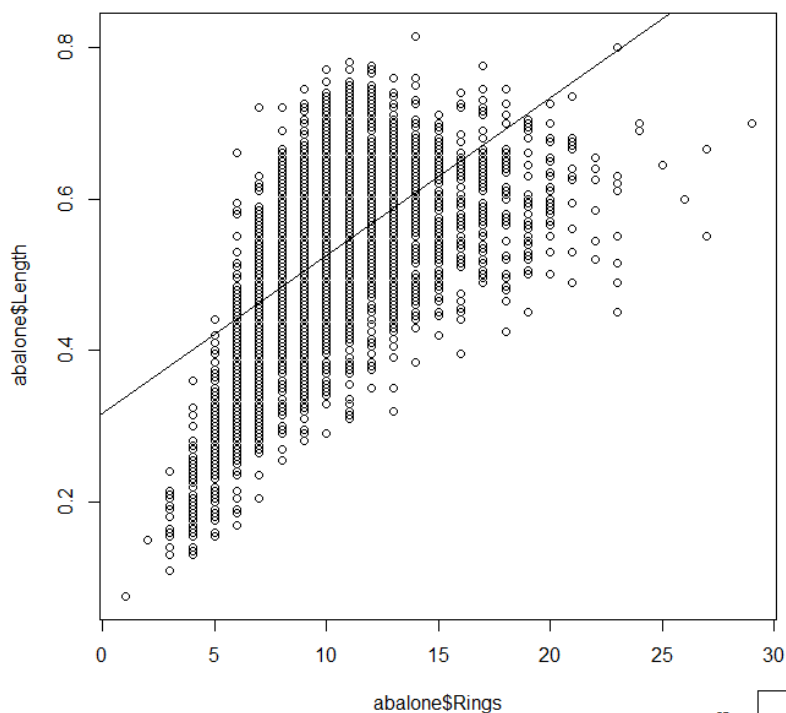


Візьмемо набір даних, що описує стать, вік та розміри молюсків галіотісів.

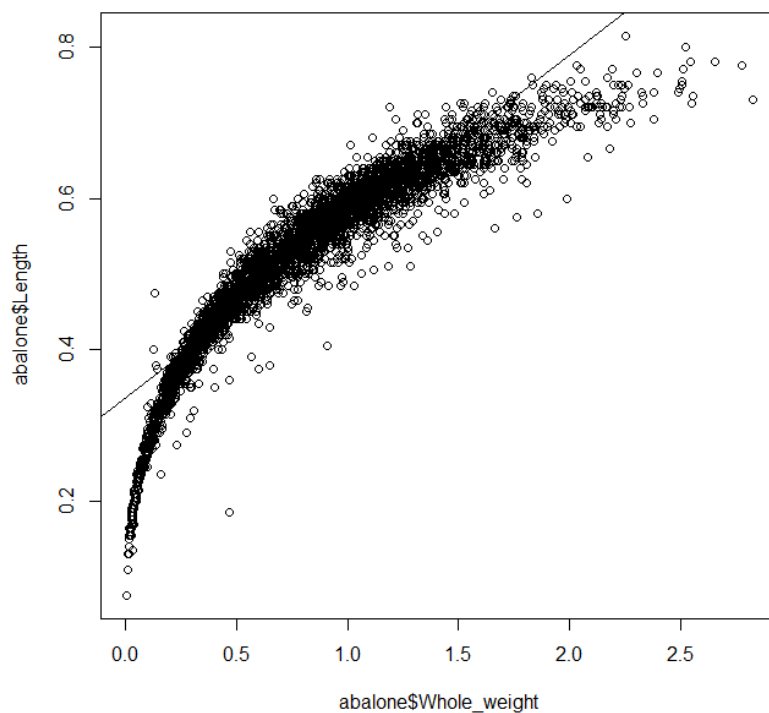
	Sex	Length	Diameter	Height	Whole_weight	Meat_weight	Gut_weight	Shell_weight	Rings
1	M	0.350	0.265	0.090	0.2255	0.0995	0.0485	0.070	7
2	F	0.530	0.420	0.135	0.6770	0.2565	0.1415	0.210	9
3	M	0.440	0.365	0.125	0.5160	0.2155	0.1140	0.155	10
4	I	0.330	0.255	0.080	0.2050	0.0895	0.0395	0.055	7
5	I	0.425	0.300	0.095	0.3515	0.1410	0.0775	0.120	8
6	F	0.530	0.415	0.150	0.7775	0.2370	0.1415	0.330	20

Діаграми, що зображують зв'язки між значеннями (в дужках коефіцієнт кореляції):



а) кількістю кілець та довжиною (0.5571226)

б) довжиною і вагою (0.9252551)



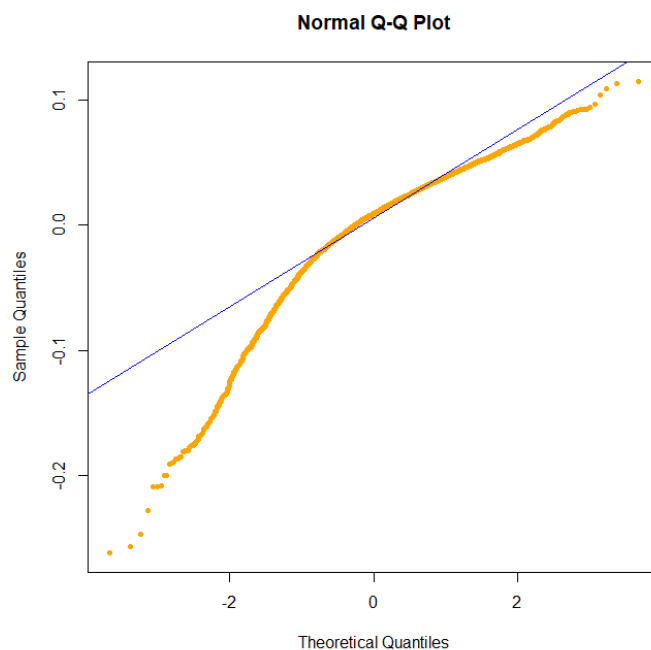
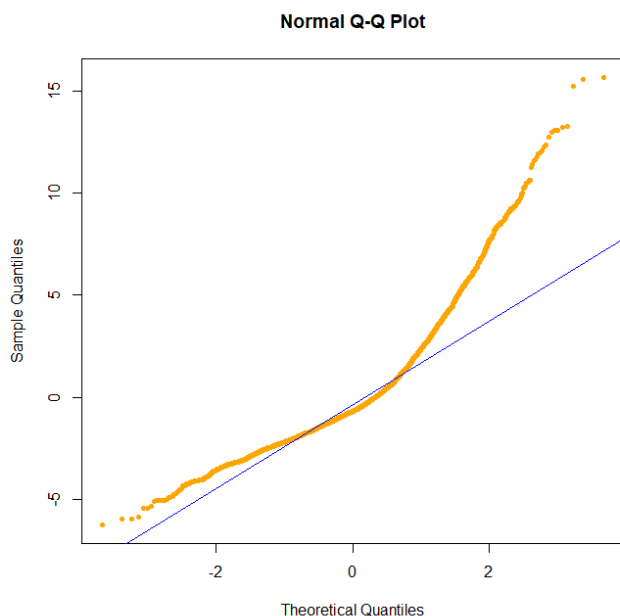
Проведемо діагностику лінійних моделей для першого і другого випадків.

У першому випадку не зовсім виконується умова лінійності та коефіцієнт кореляції досить не високий через те, що значення варіюються у певному діапазоні. Однак, разом з тим певна лінійність помітна, якщо дивитися на нижню межу значень.

У другому випадку лінійність більш очевидна і коефіцієнт кореляції чудово це відображає.

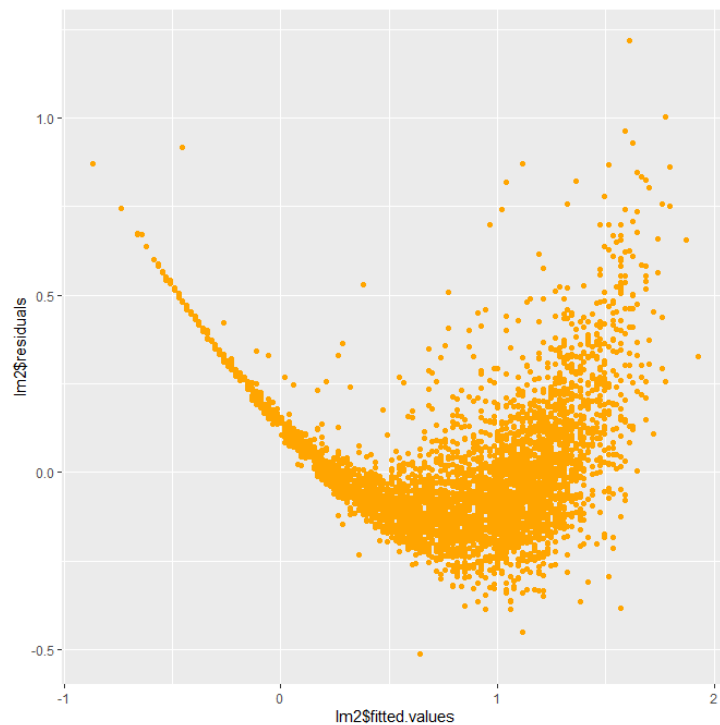
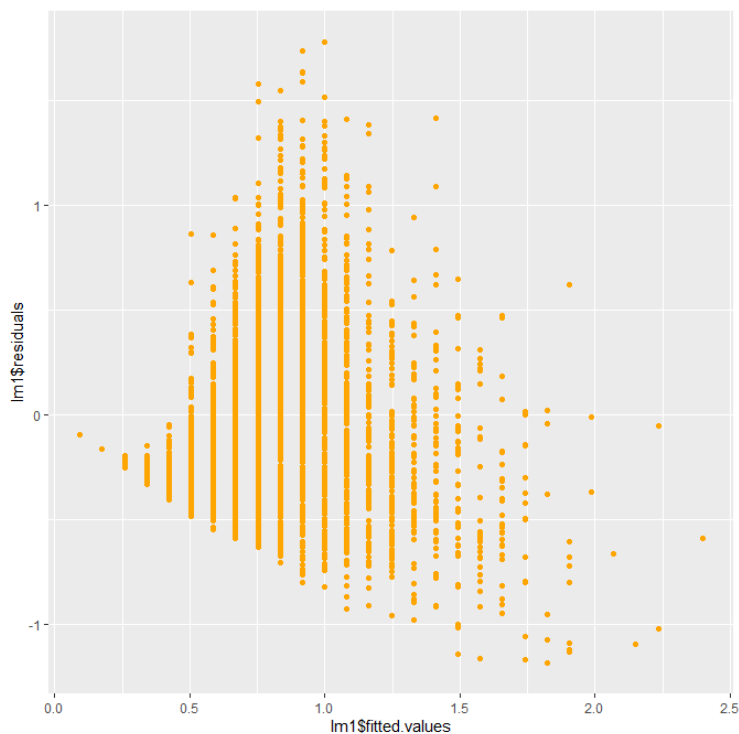
Здійснимо аналіз залишків в обох моделях.

Залишок - це різниця між реальними даними даними та даними, які відповідають моделі.



Залишки не зовсім відповідають нормальному розподілу в обох випадках, але загальна тенденція зростання відповідно до лінії нормального розподілу співпадає.

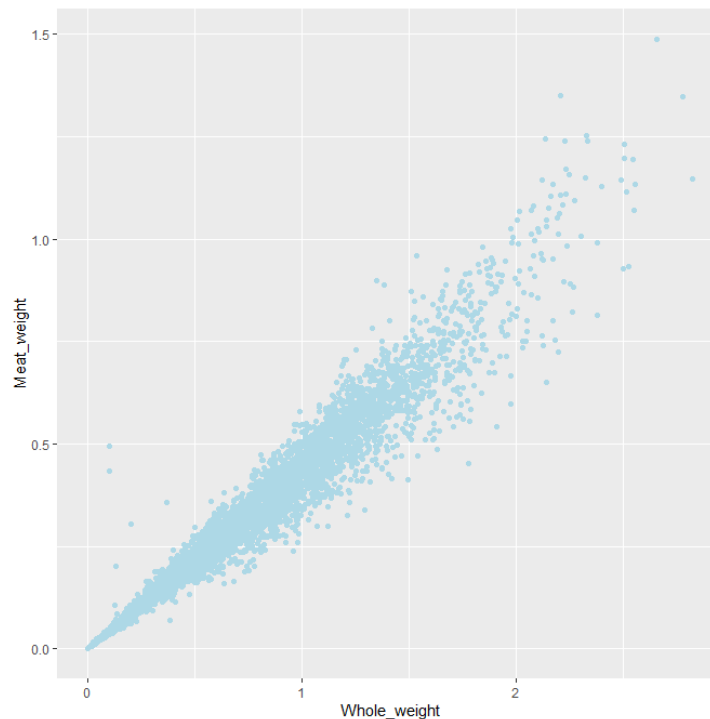
Оцінюємо варіативність залишків:



В жодному випадку умова сталості залишків не виконується.

На основі проведених досліджень, можемо стверджувати, що умови для побудови моделі лінійної регресії не виконуються для жодного набору, тобто не можна провести лінійну залежність між віком молюска галіотіса та його вагою, а також між довжиною та вагою. Можна зробити висновок, що ці дані залежні від більшої кількості умов, ніж ми досліджували.

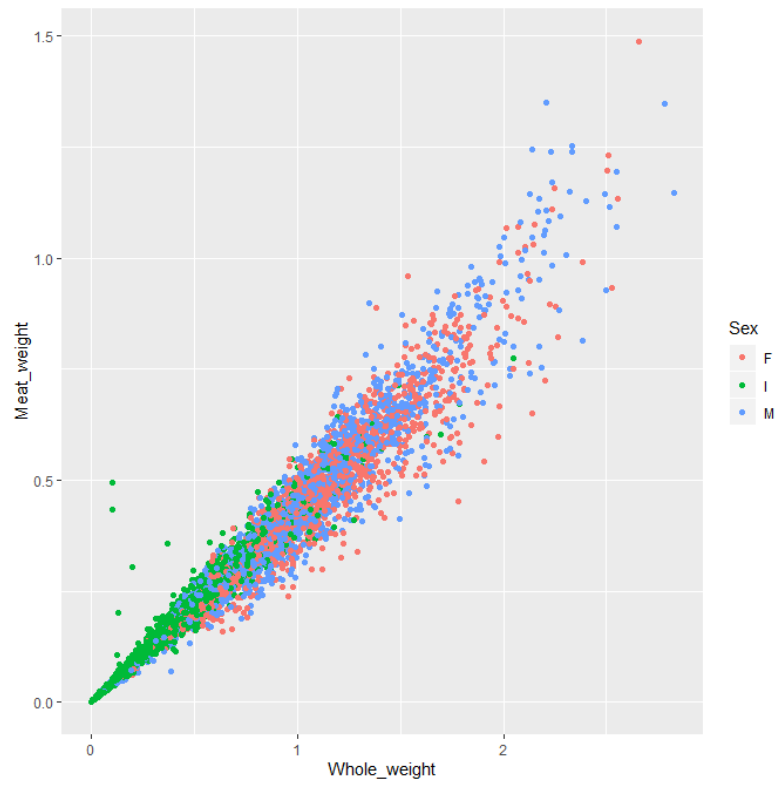
У другій частині експерименту спочатку дослідимо зв'язок між вагою м'яса молюска і його повною вагою за допомогою графіка розсіювання.



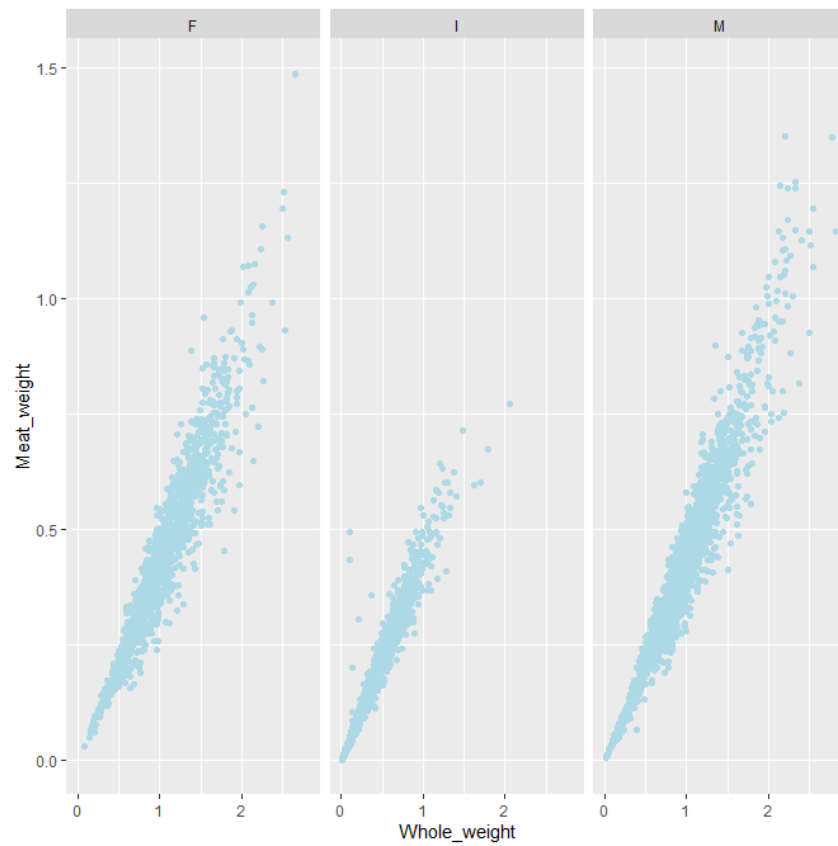
Коефіцієнт кореляції: 0.9694027

Давайте подивимось, як розподілена вага м'яса та повна вага в залежності від статі молюска.

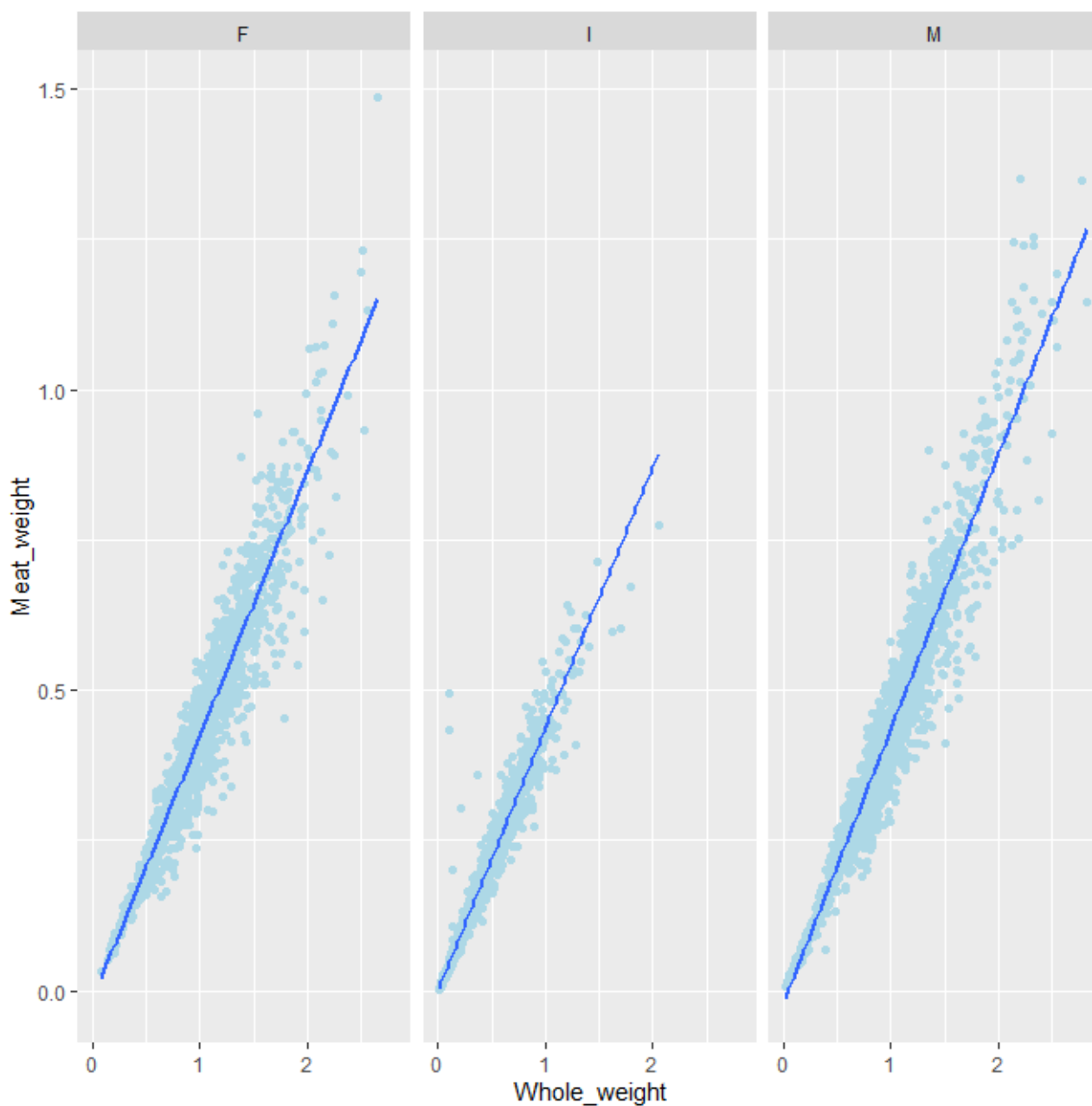
Можна побачити, що у молюсків, які ще не вирости обидві ваги менші, а у жіночої статі вага м'яса дещо менша за відповідну вагу у особин чоловічої статі. Крім того, можна побачити, що повна вага більша також у молюсків чоловічої статі.



Інший тип відображення:



Додамо до графіка ще лінію лінійної регресії:



Із цього графіка можна зробити висновок про досить високу вірогідність можливості існування лінійної залежності між повною вагою і вагою м'яса, як у дитинчат, так і у дорослих жіночих та чоловічих особин. Особливо це було б помітно, якщо очистити графіки від нестандартних значень.