**Nome:** Henrique Pablo Pinheiro dos Santos Pimentel

**Estatísticas qualitativas:**

Para começar foi realizado o calculo da porcentagem de cara variável dentro das variáveis qualitativas, como por exemplo:

Em Escolaridade foi calculado a porcentagem de Ensino médio incompleto, Ensino médio completo, Ensino superior incompleto e Ensino superior completo. Dando um resultado de:

|  |  |
| --- | --- |
| Ensino médio incompleto | 30% |
| Ensino médio completo | 20% |
| Ensino superior incompleto | 35% |
| Ensino superior completo | 15% |

Em Classe social foi calculado a porcentagem de Baixa, Média e Alta. Dando um resultado de:

|  |  |
| --- | --- |
| Baixa | 25% |
| Média | 55% |
| Alta | 20% |

Em Sexo foi calculado a porcentagem de Feminino e Masculino. Dando um resultado de:

|  |  |
| --- | --- |
| Feminino | 65% |
| Masculino | 35% |

Em Estado civil foi calculado a porcentagem de Solteiro, Namorando, Casado e Divorciado. Dando um resultado de:

|  |  |
| --- | --- |
| Solteiro | 35% |
| Namorando | 35% |
| Casado | 15% |
| Divorciado | 15% |

**Gráficos das variáveis qualitativas**

Com os resultados das porcentagens anteriores e sabendo que o gráfico de pizza é o melhor para representar dados percentuais com poucas variáveis. Foi implementado um gráfico de pizza em cada modelo. Além disso, também foi implementado um gráfico de barra fazendo a correlação dos dados de duas variáveis qualitativas, como pode ver os exemplos a baixo.

**Gráfico de Escolaridade:**

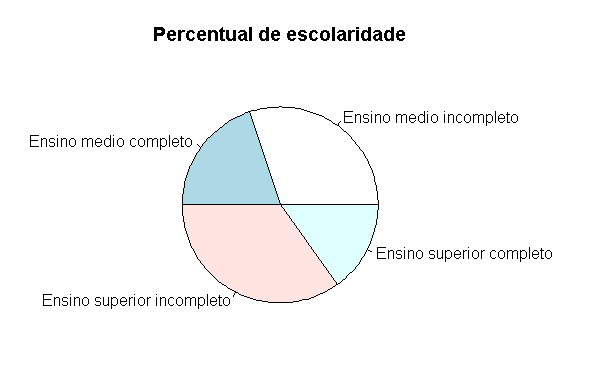


Gráfico representando que de 20 pessoas, 30% delas tem o ensino médio incompleto, 20% delas tem o ensino médio completo, 35% tem o ensino superior incompleto e 15% tem o ensino superior completo.

**Gráfico de Classe Social:**

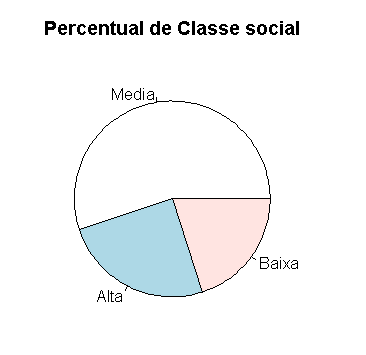


Gráfico representando que de 20 pessoas, 55% são de classe média, 25% de classe baixa, e 20% de classe alta.

**Gráfico de Sexo:**

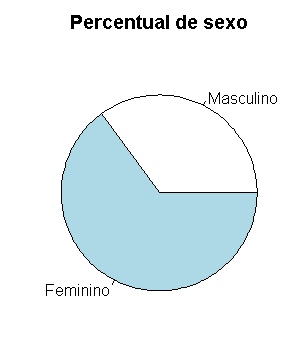


Gráfico representando que de 20 pessoas, 65% se considera do sexo Feminino e 35% se considera do sexo Masculino.

**Gráfico de Estado Civil:**

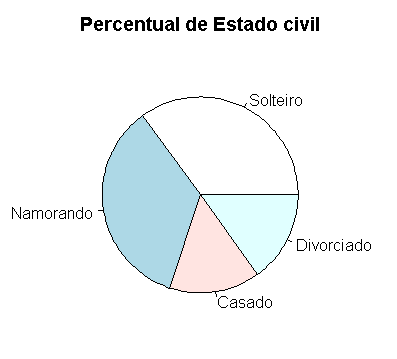


Gráfico representando que de 20 pessoas, 35% estão solteiras, 35% delas estão namorando, 15% são casadas e 15% são divorciadas.

**Gráfico de barra relacionando Classe social com Escolaridade**

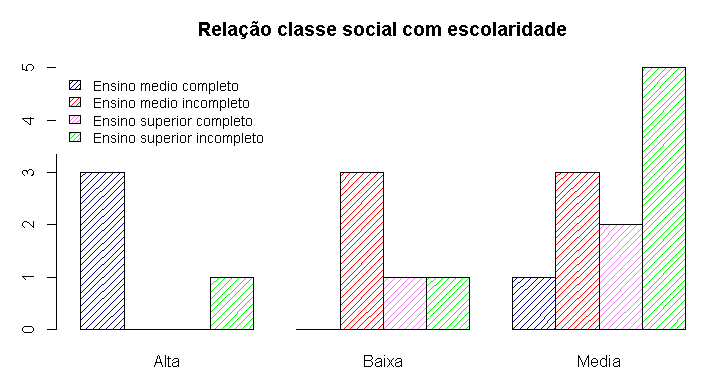


Gráfico representando um relacionamento da classe social com o nível de escolaridade, vendo que de 20 pessoas, 4 pessoas são de classe alta, 3 delas com o ensino médio completo e 1 delas com o ensino superior incompleto. 5 pessoas são de classe baixa, sendo 3 delas com o ensino médio incompleto, 1 delas com ensino superior completo e 1 delas com ensino superior incompleto. E por fim as outras 11 pessoas são de classe media, sendo 1 delas com o ensino médio completo, 3 delas com o ensino médio incompleto, 2 delas com o ensino superior completo e 5 delas com o ensino superior completo.

**Estatísticas Quantitativas:**

Para as estatísticas quantitativas foram analisados as variáveis n° de computadores, n° de animais domésticos, n° de filhos, idade, peso e altura, e foi realizada uma média, mediana, Q1, Q3, variância e desvio padrão de cada uma dessas variáveis, sendo o resultado dado por:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ####### | **Media** | **Mediana** | **Q1** | **Q3** | **Variância** | **Desvio Padrão** |
| **N° de computadores** | 0,95 | 1 | 0 | 1 | 0,6815 | 0,8255 |
| **N° de animais** | 0,75 | 1 | 0 | 1 | 0,5131 | 0,7163 |
| **N° de**  **filhos** | 0,55 | 0 | 0 | 1 | 0,4710 | 0,6863 |
| **Idade** | 24,2 | 22 | 20 | 26,25 | 38,06 | 6,1695 |
| **Peso** | 76,2 | 74 | 72 | 82,25 | 115,85 | 10,763 |
| **Altura** | 1,75 | 1,735 | 1,68 | 1,79 | 0,0074 | 0,0864 |

**Gráficos das variáveis quantitativas**

Com os resultados da tabela anterior foi feito alguns gráficos de acordo com o estilo da analise proposta em cada gráfico utilizando histogramas para revelar a distribuição de uma única variável, gráficos de dispersão para fazer uma correlação dos dados de duas variáveis, e gráfico boxplot para ver de uma forma mais detalhada os valores mínimo, médio, máximo e fora da curva de varias variáveis ao mesmo tempo.

**Gráfico Histograma de Altura**

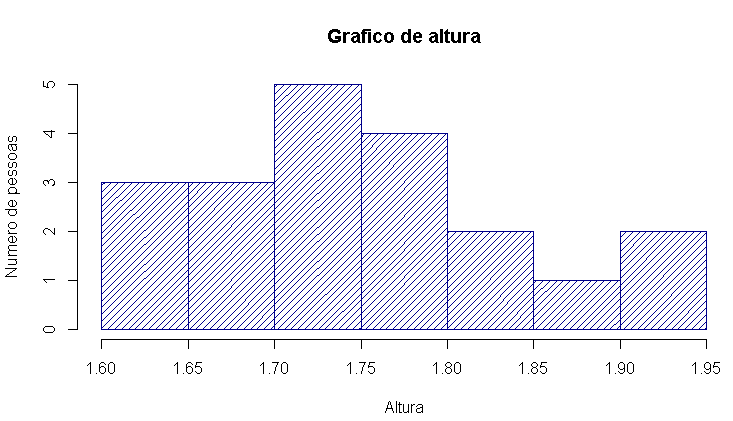
****

Gráfico de histograma representando a quantidade de pessoas dentro de uma faixa de altura representada em metros, mostrando que tem, 3 pessoas de 1,60 a 1,65, 3 pessoas de 1,65 a 1,70, 5 pessoas de 1,70 a 1,75, 4 pessoas de 1,75 a 1,80, 2 pessoas de 1,80 a 1,85, 1 pessoa de 1,85 a 1,90 e 2 pessoas de 1,90 a 1,95. Esse gráfico mostra que a media de altura dessas 20 pessoas é realmente em torno de 1,75 como demonstradas na tabela anterior.

**Gráfico Histograma de Peso**

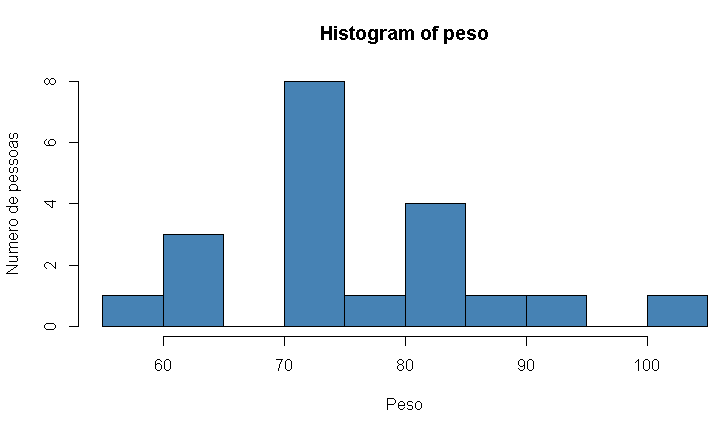
****

Gráfico de histograma representando a quantidade de pessoas dentro de uma faixa de peso representada em Kg, mostrando que tem 1 pessoa abaixo de 60kg 3 pessoas entre 60 e 65 kg, 8 pessoas entre 70 e 75 kg, 1 pessoa entre 75 e 80kg, 4 pessoas entre 80 e 85kg, 1 pessoa entre 85 e 90kg, 1 pessoa entre 90 a 95kg e 1 pessoa com mais de 100kg, entre 65 e 70kg e 95 e 100kg não tem ninguém com essa faixa de peso. Olhando no gráfico da pra ver claramente que tem mais gente na faixa entre 70 e 80kg, logo, a média, mediana e o desvio padrão da tabela a cima fazem total sentido.

**Gráfico de Dispersão de Altura com Peso**

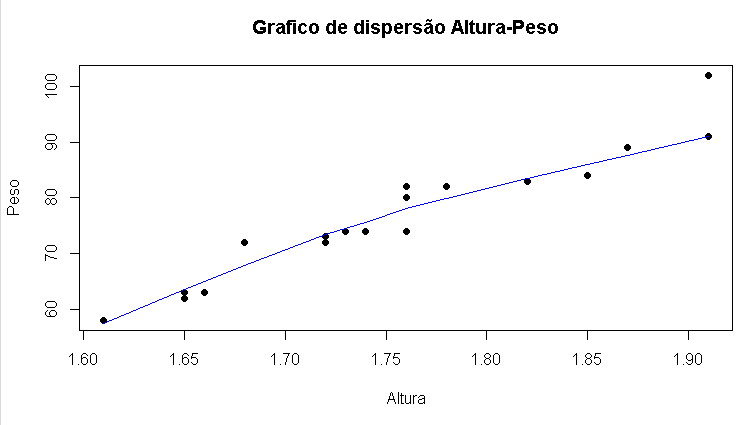
****

Gráfico de dispersão entre altura e peso, onde é possível enxergar certo padrão, onde quanto maior a altura o peso também aumenta, sendo em alguns casos variando um pouco para cima ou para baixo da linha de regressão. Logo, é possível notar um padrão como, por exemplo, as pessoas ate 72kg não passam de 1,68 que é equivalente a relação de Q1 de altura e peso. Além disso, a maior quantidade de pessoa esta concentrada entre 1,70 a 1,80 e possuem uma media de 70 a 85 kg, que acaba por ser uma boa relação entre a altura media e o peso médio, que se pode conferir na tabela vista anteriormente.

**Gráfico de Dispersão de Idade com Numero de filhos**

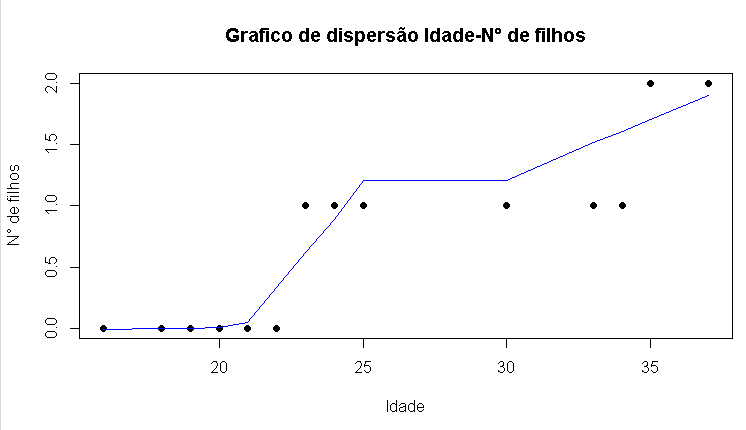
****

Gráfico de dispersão entre idade e numero de filhos, onde é possível enxergar um determinado padrão em quanto maior a idade, maiores são as chances de uma das 20 pessoas terem ao menos um filho. Nesse gráfico mostra que a baixo dos 23 anos nenhuma das pessoas possui filhos, a partir dos 23 ate em torno dos 33 anos algumas dessas pessoas tiveram no máximo 1 filho, e depois dos 33 anos foi que veio aparecer algumas pessoas com 2 filhos. Nesse caso é possível ver que dentre dessas 20 pessoas os primeiros 25% (Q1) delas é 0 enquanto os primeiros 75% (Q3) delas é no máximo 1.

**Gráfico Boxplot de N° de computadores, N° de animais de estimação e N° de filhos.**

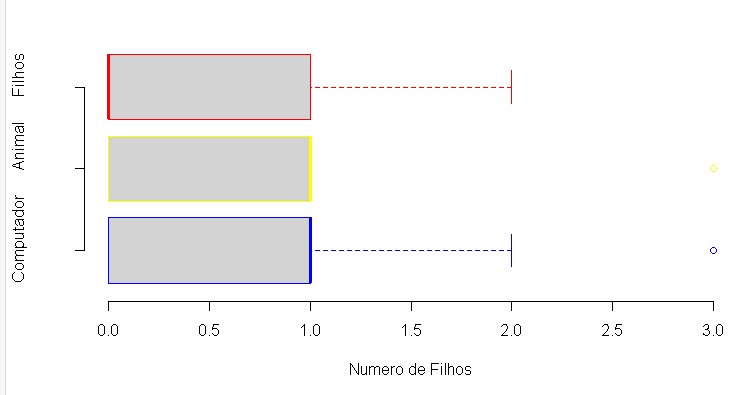
****

Gráfico de boxplot representando tanto o numero de computadores, quanto o numero de animais de estimação e o numero de filhos, a borda da esquerda representada pelo primeiro quartil (Q1), a borda da direita representada pelo terceiro quartil (Q3), a linha mais grossa representando a mediana, a linha pontilhada representando o valor máximo sem considerar os valores fora da curva, e a bolinha vazia representando os valores fora da curva, ou seja, um dado que é muito discrepante comparado aos outros. Sendo assim nesse gráfico nos conseguimos perceber que o Q0 de todas as três variáveis são 0 e o Q1 de todas as 3 variáveis são 1, mas a mediana que é representada pela linha mais grossa, no numero de filhos é 0 enquanto no numero de animais e numero de computadores é 1. Também é possível ver que o valor máximo de numero de filhos e de numero de computadores é igual a 2, enquanto o valor máximo do numero de animais de estimação é igual o Q3 e a mediana que é 1. Todos esses dados batem perfeitamente com a tabela mostrada anteriormente, mas nesse gráfico como um caso especial é possível ver que no numero de filhos tem um valor que foi bastante discrepante, ou seja, bastante fora da curva e poderia dar uma falsa interpretação dos dados do gráfico que foi o valor 3 que esta marcado como uma bolinha vazia indicando que ali tem um outlaier, um valor discrepante.