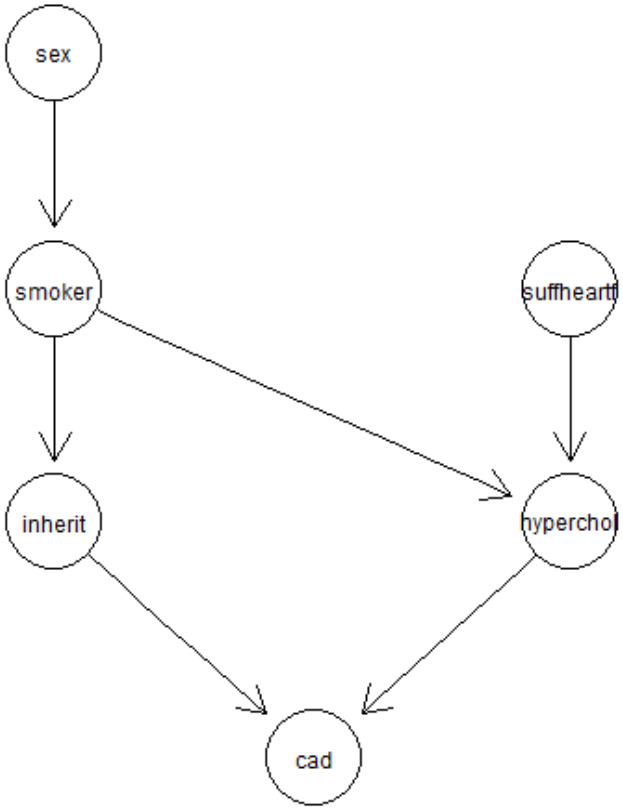
1.)

a.) Optimal Network:



d-separations in the graph:

```
> dSep(as(cad_dag, "matrix"),"sex", "suffheartf",cond = NULL)
[1] TRUE
> dSep(as(cad_dag, "matrix"),"smoker", "suffheartf",cond = NULL)
[1] TRUE
> dSep(as(cad_dag, "matrix"),"inherit", "suffheartf",cond = NULL)
[1] TRUE
> dSep(as(cad_dag, "matrix"),"smoker","cad",c("inherit","hyperchol"))
[1] TRUE
> dSep(as(cad_dag, "matrix"),"sex", "hyperchol", c("smoker", "suffheartf"))
[1] TRUE
> dSep(as(cad_dag, "matrix"), "sex", "inherit", c("smoker"))
[1] TRUE
>
```

Conditional Probability tables:

```
> s
Sex
Female   Male
    47    189
> sm.s
       Sex
Smoker Female Male
   No       17   34
   Yes      30  155
> in.sm
        Smoker
Inherit  No Yes
    No   42 120
    Yes   9  65
> sf
SuffHeartF
 No Yes
167  69
> cad.in.hyp
, , Hyperchol = No

     Inherit
CAD   No Yes
  No  69  12
  Yes 15  12

, , Hyperchol = Yes

     Inherit
CAD   No Yes
  No  35  13
  Yes 43  37

> hyp.sm.sf
, , SuffHeartF = No

          Smoker
Hyperchol No Yes
       No  27  59
       Yes 13  68

, , SuffHeartF = Yes

          Smoker
Hyperchol No Yes
       No   3  19
       Yes  8  39
```

b.) Build the network, compiled it, propagate it and then by adding the new observation the change in probabilities can be found out

Change in Probability of heart-failure and coronary artery disease (CAD):

before absorbing evidence:

```
> querygrain(grn1c, nodes = c("SuffHeartF", "CAD"), type = "joint")
     SuffHeartF
CAD          No       Yes
  No  0.3957368 0.1443930
  Yes 0.3118903 0.1479799
attr(,"class")
[1] "parray" "array"
> querygrain(grn1c, nodes = c("SuffHeartF", "CAD"), type = "conditional")
     SuffHeartF
CAD          No       Yes
  No  0.7326698 0.2673302
  Yes 0.6782138 0.3217862
attr(,"class")
[1] "parray" "array"
> querygrain(grn1c, nodes = c("SuffHeartF", "CAD"), type = "marginal")
$CAD
CAD
       No       Yes
0.5401298 0.4598702

$SuffHeartF
SuffHeartF
       No       Yes
0.7076271 0.2923729
```

---

after absorbing evidence:

```
> querygrain(grn1c.ev, nodes = c("SuffHeartF", "CAD"), type = "joint")
     SuffHeartF
CAD          No       Yes
  No  0.4078210 0.1453059
  Yes 0.2998061 0.1470670
attr(,"class")
[1] "parray" "array"
> querygrain(grn1c.ev, nodes = c("SuffHeartF", "CAD"), type = "conditional")
     SuffHeartF
CAD          No       Yes
  No  0.7373010 0.2626990
  Yes 0.6708976 0.3291024
attr(,"class")
[1] "parray" "array"
> querygrain(grn1c.ev, nodes = c("SuffHeartF", "CAD"), type = "marginal")
$CAD
CAD
       No       Yes
0.5531269 0.4468731

$SuffHeartF
SuffHeartF
       No       Yes
0.7076271 0.2923729

>
```

---

c.) Simulated a new dataset with 5 observations. Here is the new data in table format:

```
> sim.find5
     Sex Smoker Inherit CAD Hyperchol SuffHeartF
1 Female    Yes      No Yes        Yes         Yes
2 Female    Yes     Yes Yes        Yes          No
3 Female    Yes      No  No        Yes          No
4 Female    Yes      No  No         No          No
5 Female    Yes     Yes Yes         No         Yes
> |
```

Estimating the probability of “Smoker” and “CAD” given the other variables in the model:

```
$pred$Smoker
[1] "Yes" "Yes" "No"  "Yes" "Yes"

$pred$CAD
[1] "Yes" "No"  "No"  "Yes" "No"


$pEvidence
[1] 0.04488406 0.01262758 0.05544145 0.02882428 0.02075442
```

d.) Simulated a new dataset with 500 observations (data points). Saved this as “500datapoints.RData” and included it in the folder.

Estimating the probability of “Smoker” and “CAD” given the other variables in the model:

```
$pEvidence
  [1] 0.044884063 0.005291351 0.044884063 0.028824283 0.019845804 0.012627575 0.055441449 0.055441449 0.055441449 0.044884063 0.044884063 0.044884063 0.055441449
 [14] 0.005291351 0.055441449 0.055441449 0.055441449 0.012627575 0.005291351 0.055441449 0.028824283 0.019845804 0.055441449 0.055441449 0.019845804 0.055441449
 [27] 0.055441449 0.044884063 0.012627575 0.011483593 0.019845804 0.044884063 0.028824283 0.044884063 0.055441449 0.028824283 0.055441449 0.044884063 0.044884063
 [40] 0.019845804 0.028824283 0.044884063 0.019845804 0.028824283 0.055441449 0.055441449 0.019845804 0.019845804 0.028824283 0.055441449 0.019845804 0.055441449
 [53] 0.044884063 0.020754423 0.028824283 0.028824283 0.055441449 0.044884063 0.020754423 0.055441449 0.044884063 0.028824283 0.012627575 0.044884063 0.055441449
 [66] 0.020754423 0.055441449 0.019845804 0.044884063 0.044884063 0.011483593 0.055441449 0.020754423 0.019845804 0.055441449 0.055441449 0.044884063 0.044884063
 [79] 0.011483593 0.055441449 0.028824283 0.020754423 0.044884063 0.055441449 0.020754423 0.055441449 0.020754423 0.055441449 0.044884063 0.044884063 0.012627575
 [92] 0.044884063 0.044884063 0.044884063 0.028824283 0.055441449 0.044884063 0.044884063 0.055441449 0.020754423 0.055441449 0.055441449 0.028824283 0.011483593
[105] 0.044884063 0.005291351 0.044884063 0.028824283 0.020754423 0.020754423 0.055441449 0.055441449 0.055441449 0.028824283 0.019845804 0.044884063 0.028824283
[118] 0.019845804 0.055441449 0.044884063 0.005291351 0.044884063 0.012627575 0.012627575 0.055441449 0.028824283 0.028824283 0.055441449 0.028824283 0.020754423
[131] 0.044884063 0.055441449 0.019845804 0.020754423 0.011483593 0.012627575 0.011483593 0.020754423 0.028824283 0.011483593 0.019845804 0.028824283 0.012627575
[144] 0.044884063 0.044884063 0.028824283 0.020754423 0.028824283 0.055441449 0.012627575 0.044884063 0.028824283 0.020754423 0.011483593 0.011483593 0.028824283
[157] 0.028824283 0.028824283 0.011483593 0.019845804 0.055441449 0.011483593 0.019845804 0.044884063 0.020754423 0.019845804 0.028824283 0.028824283 0.020754423
[170] 0.012627575 0.019845804 0.019845804 0.055441449 0.028824283 0.044884063 0.044884063 0.028824283 0.012627575 0.044884063 0.055441449 0.055441449 0.012627575
[183] 0.020754423 0.055441449 0.055441449 0.028824283 0.044884063 0.055441449 0.028824283 0.055441449 0.055441449 0.055441449 0.044884063 0.020754423 0.028824283
[196] 0.044884063 0.028824283 0.055441449 0.012627575 0.028824283 0.011483593 0.055441449 0.012627575 0.044884063 0.020754423 0.055441449 0.011483593 0.044884063
[209] 0.028824283 0.020754423 0.044884063 0.055441449 0.011483593 0.044884063 0.019845804 0.044884063 0.019845804 0.028824283 0.055441449 0.055441449 0.011483593
[222] 0.020754423 0.055441449 0.055441449 0.020754423 0.020754423 0.055441449 0.055441449 0.044884063 0.044884063 0.019845804 0.055441449 0.044884063 0.020754423
[235] 0.055441449 0.044884063 0.028824283 0.028824283 0.028824283 0.028824283 0.044884063 0.055441449 0.020754423 0.055441449 0.055441449 0.028824283 0.044884063
[248] 0.044884063 0.055441449 0.011483593 0.044884063 0.055441449 0.020754423 0.055441449 0.055441449 0.028824283 0.044884063 0.044884063 0.055441449 0.028824283
[261] 0.019845804 0.020754423 0.044884063 0.012627575 0.020754423 0.044884063 0.055441449 0.011483593 0.055441449 0.055441449 0.019845804 0.055441449 0.055441449
[274] 0.044884063 0.055441449 0.044884063 0.019845804 0.019845804 0.055441449 0.020754423 0.044884063 0.044884063 0.044884063 0.020754423 0.012627575 0.028824283
[287] 0.044884063 0.020754423 0.055441449 0.019845804 0.044884063 0.044884063 0.055441449 0.055441449 0.055441449 0.011483593 0.019845804 0.044884063 0.055441449
[300] 0.020754423 0.055441449 0.020754423 0.044884063 0.055441449 0.044884063 0.019845804 0.055441449 0.055441449 0.055441449 0.020754423 0.044884063
[313] 0.019845804 0.019845804 0.055441449 0.044884063 0.055441449 0.055441449 0.012627575 0.020754423 0.019845804 0.055441449 0.012627575 0.044884063 0.055441449
[326] 0.055441449 0.055441449 0.019845804 0.055441449 0.028824283 0.044884063 0.012627575 0.028824283 0.028824283 0.055441449 0.044884063 0.055441449 0.011483593
[339] 0.019845804 0.055441449 0.028824283 0.028824283 0.044884063 0.055441449 0.044884063 0.044884063 0.011483593 0.019845804 0.055441449 0.012627575 0.044884063
[352] 0.044884063 0.020754423 0.020754423 0.011483593 0.044884063 0.012627575 0.044884063 0.055441449 0.044884063 0.005291351 0.019845804 0.020754423 0.019845804
[365] 0.055441449 0.028824283 0.044884063 0.019845804 0.020754423 0.005291351 0.044884063 0.028824283 0.044884063 0.055441449 0.020754423 0.055441449 0.055441449
[378] 0.044884063 0.044884063 0.044884063 0.055441449 0.019845804 0.044884063 0.055441449 0.028824283 0.055441449 0.012627575 0.028824283 0.044884063 0.019845804
[391] 0.044884063 0.028824283 0.055441449 0.020754423 0.012627575 0.055441449 0.012627575 0.019845804 0.012627575 0.044884063 0.019845804 0.028824283 0.055441449
[404] 0.019845804 0.019845804 0.055441449 0.028824283 0.028824283 0.028824283 0.012627575 0.044884063 0.012627575 0.044884063 0.012627575 0.055441449 0.019845804
[417] 0.028824283 0.055441449 0.012627575 0.055441449 0.020754423 0.044884063 0.044884063 0.020754423 0.044884063 0.019845804 0.019845804 0.055441449 0.055441449
[430] 0.044884063 0.019845804 0.055441449 0.028824283 0.028824283 0.011483593 0.044884063 0.028824283 0.055441449 0.055441449 0.028824283 0.055441449 0.019845804
[443] 0.055441449 0.019845804 0.044884063 0.044884063 0.044884063 0.055441449 0.044884063 0.019845804 0.019845804 0.055441449 0.012627575 0.012627575 0.011483593
[456] 0.012627575 0.028824283 0.044884063 0.005291351 0.044884063 0.012627575 0.055441449 0.028824283 0.044884063 0.020754423 0.019845804 0.044884063 0.055441449
[469] 0.028824283 0.044884063 0.019845804 0.044884063 0.028824283 0.005291351 0.044884063 0.019845804 0.012627575 0.011483593 0.044884063 0.020754423 0.055441449
[482] 0.055441449 0.020754423 0.028824283 0.028824283 0.020754423 0.020754423 0.020754423 0.028824283 0.044884063 0.019845804 0.020754423 0.028824283 0.044884063
[495] 0.012627575 0.012627575 0.005291351 0.019845804 0.055441449 0.044884063
```

Misclassification Rate:

```
> misclassification_rate_smoker
[1] 35.8
> tab2 = table(sim.find500$CAD, yhat$pred$CAD)
> misclassification_rate_CAD <- (1-sum(diag(tab2))/500)*100
> misclassification_rate_CAD
[1] 33
>
```

Inference:

By observing the misclassification rate for such a large dataset, we can infer that the network is performing well. It can be further improved by conditioning the probability of nodes properly, by forming a strong Bayesian network i.e. hierarchically ordering the nodes.